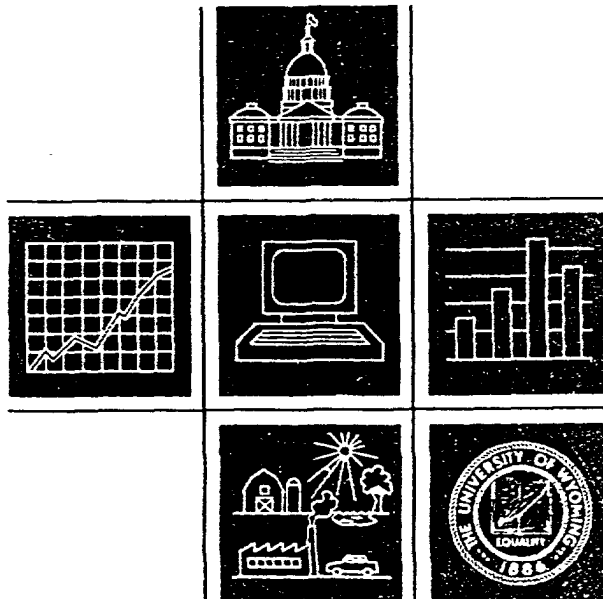
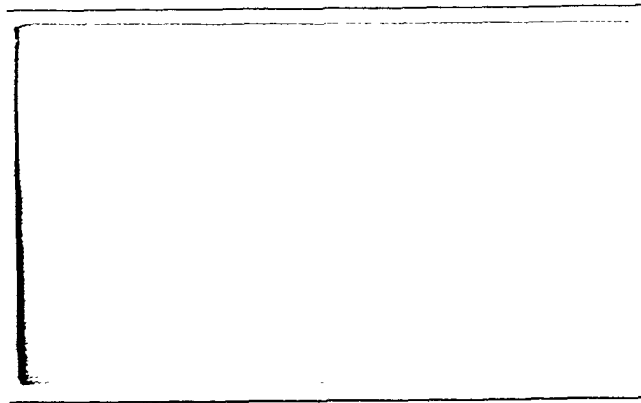


The Institute for Policy Research



... addressing public policy issues for the '80's.

VOLUME I.B

VALUING ENVIRONMENTAL GOODS: A
STATE OF THE ARTS ASSESSMENT
OF THE CONTINGENT VALUATION METHOD

EXPERIMENTAL METHODS FOR ASSESSING ENVIRONMENTAL BENEFITS

Volume I.B

Valuing Environmental Goods: A State of the Arts
Assessment of the Contingent Valuation Method

by

Ronald G. Cummings
University of New Mexico

David S. Brookshire
University of Wyoming

William D. Schulze
University of Colorado

With the Assistance of: Mark S. Walbert

Contributions by:

Kenneth Arrow
Richard Bishop
Don L. Coursey
A. Myrick Freeman
Thomas Heberlein
Daniel Kahneman
Alan Randall
Sherwin Rosen
V. Kerry Smith
Vernon Smith

USEPA Contract #CR-811077-01-0

Project Officer

Dr. Alan Carlin
Office of Policy Analysis
Office of Policy, Planning and Evaluation
U.S. Environmental Protection Agency
Washington, D.C. 20460

OFFICE OF POLICY ANALYSIS
OFFICE OF POLICY, PLANNING AND EVALUATION
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

DISCLAIMER

This report has been reviewed by the Office of Policy Analysis, Office of Policy, Planning and Evaluation, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE OF CONTENTS

<u>CHAPTER NO.</u>	<u>TITLE</u>	<u>PAGE NO.</u>
<u>PART I: THE CONTINGENT VALUATION METHOD</u>		
I.	INTRODUCTION	1
	A. The Contingent Valuation Method	1
	B. The CVM Has Achieved Acceptability...	2
	C. ... But on the Other Hand	3
	D. The Need for a State of the Arts Assessments of the CVM	5
	ENDNOTES -- Chapter I	8
II.	A HISTORICAL PERSPECTIVE FOR THE CVM ASSESSMENT	9
	A. Overview	9
	B. Social Welfare: What Is It and How Is It Measured?	10
	C. Benefit-Cost Analysis and the General Possibilities Theorem	11
	D. Market Prices as Measures of Social Welfare	12
	E. Valuing Public Goods	13
	F. Developing The CVM	16
	G. Related Research in Other Disciplines	18
	H. The Structure for a CVM Assessment	21
	ENDNOTES -- Chapter Two	22
III.	APPLICATIONS OF THE CVM: AN OVERVIEW OF ISSUES	24
	A. Overview	24
	B. Strategic Bias and the CVM	25
	C. Axiomatic Behavior and CVM Valuations	32
	D. Bias Issues Related to the Design of CVM Questionnaires	34
	1. Starting Point Bias	34
	2. Vehicle Bias	37
	3. Information Bias	38
	E. Willingness to Pay vs. Willingness to Accept	41
	F. Concluding Remarks	43
IV.	EXPERIMENTAL ECONOMICS: IMPLICATIONS FOR THE CVM	44
	A. Introduction	44
	B. Methodological Developments in Experimental Economics	46
	C. Recent Applications of Laboratory Methods Related to CVM Developments	50
	D. Valuations Under Uncertainty Conditions: Relevant Results From Laboratory Experiments	53

	E. An Example: Revelation of Compensating Income Variation	55
	1. Solicited Compensating Variations	55
	2. Tatonnement Version of the Second-Price Auction	57
	F. Conclusions	59
ENDNOTES	-- Chapter Four	60
	IMPUTING ACTUAL BEHAVIOR FROM CHOICES MADE UNDER HYPOTHETICAL CIRCUMSTANCES	61
	A. The Issues	61
	B. Hypothetical Payment: An Incentive for Accuracy	63
	C. Hypothetical Biases Related to Time	66
	D. Perceptions, Framing and the CVM	70
	1. Perceptions	70
	2. Framing	74
	E. Attitudes vs. Intended Behavior	80
	F. Concluding Remarks	82
ENDNOTES	-- Chapter V	85
VI.	COMPARISON STUDIES: WHAT IS ACCURACY?	88
	A. Introduction	88
	B. Value Comparisons: The CVM and the TCM	
	1. Knetsch and Davis	90
	2. Bishop and Heberlein	91
	3. Desvousges, Smith and McGivney	93
	4. Seller, Stoll and Chavas	98
	5. Thayer	108
	6. Fisher	109
	C. Value Comparisons: The CVM and the HPM	111
	1. Brookshire <u>et al.</u>	111
	2. Cummings <u>et al.</u>	112
	3. Brookshire <u>et al.</u> (1984)	115
	D. What is Accuracy?	118
	E. Implications for Assessments of the CVM	127
	F. Final Remarks	133
ENDNOTES	-- Chapter VI	137
 <u>PART II: THE ASSESSMENT CONFERENCE AND CONCLUSIONS AS TO THE STATE OF THE ARTS OF THE CVM</u>		
VII.	THE ASSESSMENT CONFERENCE: OVERVIEW	138
VIII.	THE POSSIBILITY OF SATISFACTORY BENEFIT ESTIMATION WITH CONTINGENT MARKETS --- BY ALAN RANDALL	139
	A. A Theory of CVM-Participant Behavior	140
	1. Value Formation	140
	2. Value Reporting	141
	3. Implications	143

	B.	"Information Bias" and Policy Evaluation	145
	C.	Concluding Comments	146
ENDNOTES	--	Chapter VIII	147
IX.		DOES CONTINGENT VALUATION WORK?	
	---	BY RICHARD C. BISHOP AND THOMAS A HEBERLEIN	149
	A.	Experimental Results	151
	1.	Goose Study Design	151
	2.	Goose Study Analysis	151
	3.	Goose Study Interpretation	153
	4.	Deer Study Design	156
	5.	Preliminary Results	157
	6.	Deer Study Interpretation and Plan for Further Research	760
	B.	Implications for CBS	163
	1.	Is Contingent Valuation Biased?	163
	2.	Do CBS Deal Adequately with Accuracy Issues?	164
	3.	Does Iterative Bidding Improve Accuracy?	165
	4.	Are Experimental Approaches the Key to Assessing and Improving CVM?	166
	5.	Should Willingness to Accept be Abandoned in CVM Applications?	168
	6.	Will the Application of Vickery Auctions Improve the Accuracy of CVM?	169
	7.	Is Attitude-Behavior Research Relevant to CVM Research?	171
	C.	Our Assessment of CVM	177
ENDNOTES	--	Chapter IX	179
X.		ON ASSESSING THE STATE OF THE ARTS OF THE CONTINGENT VALUATION METHOD OF VALUING ENVIRONMENTAL CHANGES	
	---	BY A. MYRICK FREEMAN III	180
	A.	Introduction	180
	B.	Criteria	182
	C.	Bias	185
	1.	Strategic Bias	185
	2.	Starting Point Bias	186
	3.	Information Bias	186
	4.	Vehicle Bias	187
	5.	Summary	187
	D.	Random Error	188
	E.	Assessment	191
	F.	Two Miscellaneous Comments	193
	G.	Conclusions	194
ENDNOTES	--	Chapter X	195
XI.		TO KEEP OR TOSS THE CONTINGENT VALUATION METHOD	
	---	BY V. KERRY SMITH	196
	A.	Introduction	196
	B.	Non-CVM Data: How Objective Are They?	199

	C. Tasks Requested of Survey Respondents	207
	D. The Problems with Hypothetical Questions	209
	E. The Bottom Line	212
ENDNOTES	-- Chapter XI	214
APPENDIX	-- Chapter XI	217
	Some Quibbles on the Cummings, Brookshire, Schulze -- Summary of CVM Research	217
	1. Starting Point Bias	217
	2. Iterative Bidding	217
	3. The Desvousges, Smith, McGivney Comparative Analysis	217
XII.	THE REVIEW PANEL'S ASSESSMENT	219
	A. Introduction	219
	B. Comments by Professor Kenneth Arrow	220
	C. Comments by Professor Daniel Kahneman	226
	1. ROC #5: The CVM should only be used for problems that have a "purchase structure".	226
	2. ROC #6: The use of CVM should be restricted to user values, rather than to ideological values.	231
	3. ROC #7: Accurate description of payment mode is essential to the CVM.	235
	D. Comments by Professor Sherwin Rosen	236
	E. Comments by Professor Vernon Smith	240
XIII.	SUMMARY AND CONCLUSIONS	248
	A. Overview	248
	B. Structural Biases in the CVM	249
	1. Strategic Bias	249
	2. Starting Point Bias	249
	3. Information Bias	250
	4. Vehicle Bias	252
	5. Conclusions	252
	C. Hypothetical Biases in the CVM	254
	1. Preference Research Issues	254
	2. The Comparability of WTP and WTA Measures	260
	3. Attitude vs. Intended Behavior	264
	4. Hypothetical Biases in the CVM: Conclusions	266
	D. The Accuracy of CVM Measures of Value	270
	1. Overview of the "accuracy" issue	270
	2. What is Accuracy?	271
	3. Reference Accuracy and Public Good Values	273
	4. The Need for Accuracy or Calibration Research	273
	5. The Reference Operating Conditions	273

	E. The State of the Arts of the Contingent Valuation Method	278
	1. The CVM Without Apology	278
	2. Conclusions Concerning Accuracy	279
	3. Final Remarks	280
	F. Critical Issues for Future CVM Research	282
APPENDIX	-- Chapter XIII	284
	Some Comments on the State of the Arts Assessment of the Contingent Valuation Method Draft Report --- Robert Cameron Mitchell and --- Richard T. Carson	284
	A. Aspects of the Use of the Bidding Game Format (Chapter III)	284
	1. Starting Point Bias	284
	2. Iterative Format	285
	3. Payment Card	285
	B. The Auction Analogy (Chapter IV)	287
	1. Second Price Auction	287
	C. Utility of Laboratory Experiments	288
	D. The Effect of Hypotheticality (Chapter V)	289
	1. Hypothetical Bias	289
	2. Tests of Whether Actual vs. Hypothetical Payment Makes a Difference	289
	E. Political Markets and Reference Operating Conditions (Chapter VI)	292
	F. On the Question of Accuracy (Chapter VI)	294
ENDNOTES	-- Appendix to Chapter XIII	295
REFERENCES		297

LIST OF TABLES

<u>TABLE NO.</u>	<u>TITLE</u>	<u>PAGE NO.</u>
3.1	Iterative Bidding and the Payment Card Approach	36
3.2	Measures of WTP and WTA	42
6.1	Summary of Results	92
6.2	Comparison of Benefit Estimates for Water Quality Improvements	95
6.3	A Comparison of Contingent Valuation and Generalized Travel Cost Benefit Estimates	97
6.4	A Comparison of Contingent Valuation and Contingent Ranking Benefit Estimates	99
6.5	Results of the TCM	102
6.6	Results of the Open-Ended CVM	104
6.7	Results of the Close-Ended CVM	106
6.8	Logit Analysis of the Close-Ended Form of the Contingent Valuation Method	107
6.9	Bidding Game and Site Substitution Results	109
6.10	Tests of Hypotheses from the BTSD Study	113
6.11	Elasticity Measures	115
6.12	Summary of Results from Comparison Studies	125
6.13	Reference Operating Conditions and Implied Reference Accuracy	130
6.14	ROC's Satisfied in Selected Applications of the CVM	131
9.1	Regression Analysis of Simulated and Contingent Markets for Willingness-to-Sell Goose Hunting Permits	153
9.2	Willingness-to-Accept Compensation for Sandhill Deer Hunting Permits	158
9.3	Willingness-to-Pay for Sandhill Deer Hunting Permits	159
11.1	Selected Sample of Sources for Non-CVM Economic Data	200

LIST OF FIGURES

<u>FIGURE NO.</u>	<u>TITLE</u>	<u>PAGE NO.</u>
3.1	Examples of Bid Distributions	29
4.1	Overall Average Experimental Responses	51
4.2	Group Willingness-to-Pay Function	56
4.3	Vickery Auction of N Units	58
6.1	Order of Magnitude Estimates	122
6.2	ROC's and Market, Non-Market Commodities	135
9.1	Schematic of Attitude-Behavior Relationships	172
12.1	A Hypothetical Value Function	228
12.2	Expressed Willingness-to-Pay Tax for Cleanup to Preserve Fishing in Muskoka, Haliburton and all Ontario	232
12.3	WTA - WTP Relationships	242
12.4	Tentative Results from University of Arizona Experiments	244
12.5	Change in Seller Surplus Relative to Total Surplus	246

PART I

THE CONTINGENT VALUATION

METHOD

I. INTRODUCTION

A. THE CONTINGENT VALUATION MEHTOD

The purpose of this book is that of assessing the state of the arts of the contingent valuation method (CVM) as this method is used to estimate values for public goods in general, and for environmental goods in particular. The CVM is a survey method, the essence of which is succinctly expressed by Randall et al. (1983) as follows:

"Contingent valuation devices involve asking individuals, in survey or experimental settings, to reveal their personal valuations of increments (or decrements) in unpriced goods by using contingent markets. These markets define the good or amenity of interest, the status quo level of provision and the offered increment or decrement therein, the institutional structure under which the good is to be provided, the method of payment, and (implicitly or explicitly) the decision rule which determines whether to implement the offered program. Contingent markets are highly structured to confront respondents with a well-defined situation and to elicit a circumstantial choice contingent upon the occurrence of the posited situation. Contingent markets elicit contingent choices." (p. 637)

The use of surveys as a means for obtaining values from individuals elicits in many a feeling of uneasiness. This may be attributable in part to the association of surveys with opinion polls and the general awareness that such polls may not be reliable: in 1948, opinion polls 'elected' Mr. Dewey, but voters elected Mr. Truman. As is discussed later, psychologists would generally support the notion that opinion polls may be unreliable; their research demonstrates that opinions, or attitudes, may be poor predictors of actual behavior.

In the CVM, however, individuals are asked neither about their opinions nor about their attitudes: they are asked about their contingent valuation (if 'this' happens, what would you be willing to pay). However, while questions posed in the CVM are (arguably) not attitudinal, the 'market', the commodity and the payment, as they appear in the CVM, are hypothetical. As will be seen, a large part of the criticisms of the CVM in terms of the reliability or accuracy of value measures drawn therefrom, arise from the hypothetical nature of the CVM.

The CVM has strengths and it has weaknesses. Experimental efforts to develop the method -- devise ways to mitigate or eliminate weaknesses and enhance strengths -- began but a decade ago; prior to 1978, only a handful of scholars were involved in its development. As interest in applications of the CVM increased, and its presence became more broadly recognized in the research community, more and more scholars have entered the debate as to the efficacy of the CVM, in real and potential terms, as a means for valuing public goods. At this point in time, a substantial literature has developed concerning the issue, in the most general terms, as to whether or not one can hope to derive meaningful measures of individual values from a method wherein all aspects 'relevant' to value decisions are artificial, or hypothetical. A brief overview of this literature will provide the reader with some flavor for this controversy and, therefore, with an appreciation for the major objectives of this book -- a topic which will be discussed below. Thus, in the following two sections we consider arguments related to the proposition: The CVM has achieved acceptability (section B) ..., but on the other hand (section C).

B. THE CVM HAS ACHIEVED ACCEPTABILITY...

Randall and others argue that research to date has established the acceptability of the CVM as a method for non-market benefit estimation and that the current task "... is to identify and explain systematically the relationship between the structure and performance of contingent markets". (Randall et al., 1983, p. 642) Thus, Randall, et al. assert:

"At the outset, the research agenda in contingent valuation sought to establish, in the face of considerable skepticism, contingent valuation as an acceptable method of non-market benefit estimation (acceptable in the sense that it works about as well as available alternative techniques and is adaptable to at least some valuation tasks that alternative methods cannot handle). That objective has been attained. In addition, the experimental work of others has blunted traditional fears that strategic responses would inevitably dominate data sets of stated personal valuation. (p. 642)

Other authors, despite their critique of some CVM studies, suggest cautious optimism for the promise of the CVM; for example:

"(CVM studies) are a promising approach for the estimation of non-market environmental values. There has been steady progress in minimizing biases, just as there has been progress with problems in other techniques; nevertheless, we are far from being out of the woods." (Rowe and Chestnut, 1983, p. 408)

Since the relatively recent beginning of empirical experiments with the CVM, 1/ progress of sorts has undeniably, been made in the development of the CVM. As pointed out by Randall et al. (1983), bids obtained in CVM studies are generally shown to be significantly related to income, availability of substitute and complementary commodities and demographic characteristics; i.e., CV bids "...are not random numbers." (p. 639-40) Bids have been shown to be consistent with actual behavior. (Randall, et al. pp. 639-40) As is discussed in some detail in 2 later chapter Of this book, maximum willingness-to-pay measures derived from CV studies have been shown to be consistent with market-demand-based values. Within this context, a basis exists for Randall, et al.'s assertion that "several kinds of evidence generated by ... (CVM)...studies support contingent valuation methods." (p. 639)

Moreover, in a recent study by Schulze et al. (1981), selected CVM studies were reviewed to the end of assessing the extent of various biases in CVM measures. The authors conclude that "Biases do not appear to be an overriding problem" (p. 170) although the authors point out that "...to establish a precise contingent market -- the 'good' must be well-defined" (p. 170).2/

C. ...BUT ON THE OTHER HAND.

Notwithstanding the "progress" noted above, others within the economics profession, and many outside the profession, reject the above-described notion that the CVM has attained anything near the level of "acceptability" ascribed to the method. In reviewing estimation methods, including the CVM, for valuing non-market goods, Feenburg and Mills (1980) offer the dreary conclusion that "In the absence of market data, demand or willingness to pay estimation would appear to be hopeless" (p. 58). 3/ Referring specifically to survey methods such as the CVM, Feenburg and Mills seemingly presume to speak for the economics profession in offering the following conclusion.

"Economists are biased against such surveys because they believe crucial contrary-to-fact questions are unlikely to be answered accurately. People lack the incentive and ability to answer accurately questions such as, 'How much more often would you swim in lake L if ambient pollution concentrations were reduced 10%?' Most people presumably experiment and talk to others to ascertain the effect of pollution abatement on their utility-maximizing behavior. Thus, economists doubt the accuracy of survey responses regarding effects of pollution abatement." (p. 169)

Interestingly enough, the 'incentives' criticism of measures drawn from the CVM, as couched above, is inextricably related to a second criticism of the CVM, viz, biases resulting from strategic behavior on the part of survey participants. Essentially, the strategic behavior hypothesis -- discussed in detail below in Chapter II -- posits behavior by survey respondents whereby false responses are given when such responses may result in 3 gain to the individual; i.e., "...it is in the selfish interest of each person to give false signals, to pretend to have less interest in a given collective consumption activity than, he generally has..." (Samuelson, 1954, p. 389). From empirical efforts to test the strategic behavior hypothesis, it is shown that the more hypothetical the question in a survey, the less the incentive for strategic behavior -- the use of hypothetical questions could be a means of avoiding biases from strategic behavior (Freeman, 1979a, pp. 97-99). Herein lies the potential dilemma: the more hypothetical the question, the less the incentives for strategic behavior but, also, the less are incentives for accurate responses.

In addition, to the above, two related sets of considerations which pose questions as to the efficacy of the CVM emanated from outside of the profession per se, viz, from the branch of psychology referred to as 'cognitive psychology'. The first of these (noted above) questions the extent to which responses derived in CV studies are expressions of attitudes as opposed to intended behavior (as is presupposed in CV studies) and 2 related controversy in the discipline of psychology concerning the extent to which attitudes are reliable predictors of behavior, (Bishop and Heberlein, 1979). A second set of considerations received from psychology which is of potential relevance for the CVM strikes at one of the most basic concepts in economic analysis: the concept of rational behavior. A number of recent studies point to stark discrepancies between actual decision-making behavior and the postulates of rationality, particularly in circumstances involving uncertainty. 4/ Arrow (1982) notes that "...these

failures in the rationality hypothesis are in fact compatible with some of the specific observations of cognitive psychologists" p.5). The 'observations' referred to by Arrow will receive considerable attention in later sections of this book. For present purposes, two of these observations from psychological research are germane. In direct contrast to expected utility theory wherein subjective probabilities based on prior information play a major role, cognitive psychologists argue that individuals, in evaluating uncertain events, tend to ignore both prior information and the quality of present evidence (Tversky and Kahneman, 1974, 1981). Secondly, also in direct contrast with the rationality precepts underlying expected utility theory, cognitive psychologists essentially argue that an individual's valuation of a commodity, along with many other commodities, is not simply dependent on the commodity set (prices, income and commodities), but on how the set is described -- different descriptions of the same commodity space may yield different values for specific commodities. (Tversky and Kahneman, 1981)

Implications of these observations for potential biases in results from CV studies are obviously a matter of some concern. For example the first issue -- excessive reaction to current information -- may imply that obtained CV values are susceptible to the influence of (often) temporary 'media events'; in terms of efforts to value environmental quality, the Three-Mile Island incident and the furor over Love Canal -- a popular media topic in 1980 -- come to mind. Moreover, the applicability of CV values obtained in one 'current information' climate to values relevant for a different climate is questionable. The second issue -- the dependence of commodity values on how commodities are described -- implies potential biases arising from the framing of willingness-to-pay questions in the CV questionnaire; thus, for any given public/environmental commodity to be valued via the CVM, different descriptions of the same basic commodity could yield different estimates of values of the commodity. 5/

D. THE NEED FOR A STATE OF THE ARTS ASSESSMENT OF THE CVM

It is important that the reader understand the context for which the controversy described above is relevant. President Reagan's Executive Order 12,291 (46 Fed. Reg. 13, 193, Feb., 17, 1981) requires that federal agencies such as the EPA consider the benefits and costs of federal regulations/actions prior to their implementation. For EPA regulations, such as air and/or water quality standards and regulations on hazardous waste disposal practices, cost may be amenable to estimation but benefits attributable to a large part of these regulations are non-market, 'public goods' in nature: cleaner air and water, a safer environment. Agencies such as the EPA then have strong incentives and interests in identifying and developing means by which benefits attributable to public goods -- such as environmental improvements -- may be assessed.

Methods other than the CVM exist for valuing public goods, primarily the Travel Cost Method (TCM) 6/ and the Hedonic Price Method (HPM) 7/ The environmental (and other public good) 'commodities' for which the TCM or HPM might be used for valuation purposes are very limited, however. 8/ For the broad range of air quality and environmental safety issues Of potential regulatory concern to the EPA, the CVM is, metaphorically, the only game in town for estimating relevant benefits. Obviously, the fact that the CVM is no worse than other methods or is the only game in town is not a sufficient reason for the use of CVM values as 'acceptable' economic measures of social benefits in policy assessments. However, one sees rationales like these suggested as justifications for the continued development of the method. For example, Burness et al. (1983) conclude their discussion of caveats relevant for reported CVM results with the observation:

"Continued interest and research in this (the CVM) area are clearly warranted given, first, the importance of the public goods issue and, second, the lack of apparent alternatives to some form of the survey method in deriving valuations for large classes of public (environmental) goods." (p. 682)

On the other hand, the fact that the CVM is 'the only game in town' for providing information of relevance to critical policy issues of the day is a powerful incentive for scholars to meet the intellectual challenge to devise means by which the CVM (or other methods) can be made effectual in responding to society's needs.

Within this milieu (Chapter II traces the character of historical efforts to develop the CVM), it seems fair to say that all scholars -- whatever their predilection towards the CVM -- who are directly or indirectly involved with the method appreciate the immediate need for reflective pause in CVM experiment/application activities. Such a pause is required for thinking through the many (again, intuitive) propositions that have been posed as indicative of sources for bias in CV measures, as well as related (again, often intuitive) counter-arguments. Most importantly, a reflective pause is required for a re-examination of means by which we can effectively apply the scientific method in our efforts to assess the CVM. In this regard, Joan Robinson's (1962) polemic concerning the difficulty in social sciences of applying the scientific method, is relevant for our discussions:

"(Referring to why economics is a branch of theology) The process of

science ... consists in trying to disprove theories ... The great difficulty in social sciences ... of applying scientific method, is that we have not yet established an agreed standard for the disproof of an hypothesis" (pp. 22-3) (theories become religions in the social sciences because) "first, the subject matter has much greater political and ideological content, so that other loyalties are ... involved .. (and secondly) it has been sometimes remarked that economists are more queazy and ill-natured than other scientists. The reason is that, when a writer's personal judgment is involved in an argument, disagreement is insulting." (p. 24)

As will be seen in later discussions, it is not rare to find one writer questioning the judgement of other writers in the CVM literature and there exists considerable disagreement, if not confusion, as to standards for proving or disproving hypotheses relevant for important aspects of the method. Thus, developments with the CVM have reached an important watershed at which a state of the arts assessment of the method is timely. The purpose of this book is to provide such an assessment.

The critical assessment of the literature relevant for the CVM is the substance of the remaining five chapters in Part I of this book. Given that the intent of this literature review is to go beyond a simple description of literature to an assessment of the strengths and weaknesses of the CVM, we begin in Chapter II with the development of an historical setting for the CVM within which an assessment framework for evaluating the state of the arts of the method is promulgated. Arguments developed in Chapter II will set the stage for the central thrust of remaining chapters in Part I.

The arguments developed by the authors in these five chapters are intended to serve as a point of departure for a critical examination of the state of the arts for the CVM. Obviously, the authors' assessment of the CVM is in no way "the profession's" assessment and, as noted above, what is needed at this point in time is a profession-wide evaluation of the CVM. An effort to obtain something akin to a broader, profession-wide assessment is accomplished via an Assessment Conference, which has the following form.

A "Conference on Valuing Environmental Improvements: A STATE OF THE ARTS ASSESSMENT OF THE CVM" was held in Palo Alto, California, on July 2, 1984. The purpose of the Conference was to elicit a Review Panel's judgements as to the promise of the CVM as a means for valuing public/environmental goods. The Panel consisted of leading scholars in the economics and psychology professions and included:

Kenneth Arrow, Stanford University

Daniel Kahneman, University of British Columbia

Sherwin Rosen, University of Chicago

Vernon Smith, University of Arizona

The Review Panel's consideration of the CVM was based, in addition to their general knowledge and expertise in the science of public goods valuation, upon two sets of information. The first information set was the authors' critical assessment of the CVM as set out in Part I of this book;

Part I was made available to Panel members well in advance of the Conference. The second information set was papers and presentations provided by four leading scholars involved in research related to the CVM. Papers/presentations by these scholars focused first on their critical assessment of Part I of this book and secondly on their individual assessments of the promise, strengths, and weaknesses of the CVM. The four scholars offering presentations at the Conference were:

Richard Bishop, University of Wisconsin

A. Myrick Freeman, Bowdoin College

Alan Randall, University of Kentucky

V. Kerry Smith, Vanderbilt University

Results from the conference are reported in Part II of this book. The authors' assessment of the CVM -- the substance of Part I -- and a more general, profession-wide assessment of the CVM -- Part II of the book -- allow us to conclude with what the authors hope will be regarded as an objective, benchmark evaluation of the CVM. Drawing from the diverse sources described above, in Chapter XIII the authors will offer final conclusions as to the current state of the arts for the CVM.

ENDNOTES

Chapter 1

- 1) As examples, see Davis (1963) and Bohm (1971).
- 2) These conclusions are challenged, however, in Rowe (1983).
- 3) As part of the authors' context for the cited conclusion, the authors also assert that "... almost no empirical work has been based on careful theoretical analysis" (p. 3). Excepting the use of surveys, this conclusion is softened somewhat in their Chapter 10, however.
- 4) As examples, see S. Lichtenstein and P. Slovic (1971); D. Grether and C. Plott, (1979); Kunreuther, et al. (1978); H. Simon, (1979).
- 5) For related discussions, see M.C. Weinstein and R.J. Quinn, (1983). Furthermore, it may be tempting to set this source of bias aside as one which can be readily eliminated through questionnaire design or accounted for by administering various questionnaires with alternative question frames. A careful consideration of the example given in Arrow (1982, p. 7) belies the ease by which this problem may be mitigated by questionnaire design or administration.
- 6). See R. Mendelsohn and G.M. Brown, Jr., (1983).
- 7). See S. Rosen, (1974).
- 8). See Freeman, (1979a), Chapters 4-5; particularly pp. 85-87.

II. A HISTORICAL PERSPECTIVE FOR THE CVM ASSESSMENT

A. OVERVIEW

As stated above, the CVM is a method for estimating values attributable to non-market, or public, goods. The intent of this chapter is to provide the reader with some flavor for how and why interest in the CVM was initiated, the rationale for and nature of early experimental efforts to develop the method and the evolution of our current understanding of the strengths and weaknesses imputed to the method. These discussions then serve to define the necessary scope of our inquiry as to the state of the arts of the CVM.

In establishing an historical perspective for an assessment of the CVM, we must begin by recognizing the ultimate ends sought in applications of the method. As noted above in section I.D, the need for benefit measures arises from the need for benefit-cost assessments related to environmental (more broadly, public) goods/commodities--commodities which are 'public good' in nature; of course, market prices (and their use in deriving measures for consumer surplus) are not available for such goods. Implicitly, market prices are appropriate measures of the 'benefits' (social welfare) of concern in benefit-cost assessments and, therefore, represent a standard for accuracy, or 'appropriateness', against which CVM measures are often compared.

Our historical perspective must therefore begin with a consideration of the benefit-cost analysis (BCA) framework per se in terms of its efficacy as a structure for processing information in ways that are meaningfully reflective of social welfare consequences associated with social actions; this topic is considered in sections B and C, below. In section D, we then consider the extent to which market prices, as they are commonly used in BCA, are 'appropriate' measures of social welfare, as social welfare is implicitly defined in the BCA. We will then have established some basis (which will be later expanded) for appreciating the nature of the valuation institution -- the market -- which is (arguably) a standard for assessing measures derived by the CVM. At this point, we will be prepared to begin our inquiry as to the public goods valuation issue. In section E the general valuation issue is described. A brief review of the substance of efforts to develop the CVM is given in section F and section G briefly describes the relevant, related research in the field of psychology. The chapter concludes with section H wherein an effort is made to focus earlier discussions given in sections B-G on related questions as to the necessary scope and structure of a comprehensive assessment of the state of the arts of the CVM.

B. SOCIAL WELFARE: WHAT IS IT AND HOW IS IT MEASURED?

Economists have long been concerned with questions concerning how one might define and measure economic, or social 'welfare'. 1/ In early years, a good deal of this concern focused on the debate as to the dependence of any notion of social welfare on value of judgements, a dependence argued by Robbins (1932) as out of place in scientific, objective analysis. Bergson's (1938) social welfare function provided the profession with a mechanism wherein the role of value judgements in welfare economics could be isolated and clarified: such 'non-economic' factors could be entered in the welfare function as variables just as we include 'economic factors' such as goods, services and factors of production.

While Bergson's economic welfare function provided a context for tracing implications that arise from any given set of value judgements, two major problems remained. First, some guide was required as to how one might define/delineate alternative sets of values which might lead to a useful social ordering of alternatives; secondly, how might we choose from among these alternative sets of values? These were the questions addressed by Arrow (1951). Based on five general conditions, including the condition that the social welfare function is not to be imposed or 'dictatorial' -- i.e., individual preferences count -- Arrow derives the renowned General Possibility Theorem which says, in essence, that one cannot structure a meaningful social welfare function without violating one or more of his five conditions -- particularly those related to 'counting' individual preferences. (Arrow, 1951, pp. 46-60) While the general relevance Of Arrow's theorem to welfare economics has been criticized, particularly in terms of its relevance to Bergson's welfare function 2/ the bulk of such criticisms has been dismissed by later analysis. 3/ The necessarily brief, and admittedly incomplete, sketch of early controversy concerning value judgements in a social welfare function given above is intended to set the stage for a theme which will recur throughout this book and which will be particularly important for efforts to suggest conclusions regarding the state of the arts for the CVM -- the task of the Assessment Conference.

This theme is set out in the form of two questions, developed below, and is framed within the context of benefit-cost analysis (BCA). This context is used given that the raison d'etre for our interest in the CVM is its use in generating estimates of value (benefits) for use in benefit-cost analysis related to the provision of public goods in general, and environmental commodities in particular (see section I.D above). The questions of interest in this regard are:

- (i) how are value judgments treated In BCA; i.e., how does use of the BCA square with the General Possibilities Theorem?
- (ii) to what extent are market prices, commonly used in applications of the BCA, 'appropriate' measures of social welfare (or 'benefits')?

C. BENEFIT-COST ANALYSIS AND THE GENERAL POSSIBILITIES THEOREM

While well-understood by most economists, it is useful to briefly review a basic inconsistency underlying BCA as it relates to the General Possibilities Theorem (GPT). The relevant issue is succinctly expressed by Dasgupta and Pearce (1978) as follows.

"From the point of view of BCA the main lesson of this discussion seems to be the following. BCA has been generally interpreted as a method of aggregating individual preferences so as to provide a basis for social choice. The Impossibility Theorem claims to show that no such aggregation is possible without introducing ethical judgements of a more specialized kind than requiring simply that individual preferences should count. The explicit introduction of ethical judgments into BCA thus appears inevitable." (p. 90)

Thus, since 3 social welfare function involves value judgements, the question becomes how such judgements are to be treated by BCA practitioners. Under the worst conditions, this question is simply begged. Under the best (and most common) conditions, economists simply rely on efficiency criteria, arguing that such things as distributional effects will either 'cancel out' or can be addressed by other means. 4/ In this case, the economist prepares the XX which follows from alternative sets of value judgements and leaves to the decision-maker the choice of 'appropriate' value judgements.

The central issue here is that, first, the idea of consumer sovereignty supposedly underlies the logic of BCA wherein 'values' (discussed below), or preference, are aggregated across consumers. But, following the GPT, such aggregation cannot occur without violating one or more of Arrow's 'reasonableness' criteria. We should note that even if such aggregation were justifiable, substantive ethical issues would attend the BCA result when interpreted as a measure of social welfare. 5/ Thus, BCA "... proceeds in a fashion which is at odds with its apparent philosophy". (Dasgupta and Pearce, 1978, p. 94) From this we conclude the following which will be relevant for later discussions: in using BCA for assessments of benefits/welfare accruing to society as a result of (e.g.1 the adoption of an environmental policy, measures used therein are appropriately assessed within a context which includes consideration of implied judgements as to the substance of 'social value'.

D. MARKET PRICES AS MEASURES OF SOCIAL WELFARE

As implied in the above, the maximization of net benefits derived via BCA is typically used for assessing a project's implications for social welfare. It is typically assumed that market prices for outputs and inputs serve, at least as a first approximation, as 'proper' measures for socially relevant benefits and costs. We will not further belabor the point that 'appropriate' prices must reflect an 'appropriate' objective (social welfare) function.^{6/} 'Proper' in this regard is generally taken to refer to the Pareto criterion.

It is generally appreciated that market prices are identical to the shadow prices implicit to Pareto Optimality under conditions which include: equality between market prices and marginal production costs; and equality between marginal production costs and the social opportunity costs of resources. (Dasgupta and Pearce, 1978, pp. 97-105) It is also generally appreciated that these two conditions are seldom, if ever, satisfied in the real world due to, among other reasons, the existence of externalities, imperfect competition in product and factor markets and unemployed resources. (Dasgupta and Pearce, 1978, pp. 105-109) In terms of the public sector, we note the unresolved controversy as to whether or not movements toward Pareto Optimality might result from marginal social cost pricing notwithstanding distortions in the private sector. ^{7/} In the end, one sees in the debate over the extent to which market prices may serve as 'adequate' proxies for Pareto-like shadow prices, our earlier-cited lament by Joan Robinson regarding the absence in the social sciences of standards by which hypotheses can be disproved; e.g., after reviewing this debate, Dasgupta and Pearce observe "The role of personal judgement is the real source of criticisms of imputed price estimates, since it would appear to lend a large element of 'subjectivity' to a discipline which purports to be objective ... (referring to market prices) ... using them for the purposes of BCA might be no less subjective." (Dasgupta and Pearce, 1978, p. 116)

From the above we may conclude the following. Given -- accepting -- Pareto efficiency as 'the' social welfare criterion for ranking and/or assessing the consequences of social actions, market prices serve, at best as weak approximations for relevant measures of social value.

E. VALUING PUBLIC GOODS

In the debate surrounding the social welfare function issue, relatively little attention was given to that class of goods which, when made available to one person, is made available to all because of joint supply and access to which cannot be denied to individuals via pricing policies, i.e., to 'public goods'. A formal inquiry as to the relationship between social welfare and levels of provision of public goods was introduced by Samuelson in 1954. Samuelson's conclusions of primary relevance for our discussions are as follows: First, one cannot hope to obtain values/measures of individual preferences for public goods by directly asking people to reveal their preferences: "One can imagine every person (being asked to reveal) ... his preferences by signalling in response to price parameters .. to questionnaires, or to other devices." (p. 389), but with such procedures, "... any one person can hope to snatch some selfish benefit in a way not possible under the self-policing competitive pricing of private goods ..." (p. 388). This observation has been interpreted as a rationale for rejecting the possible use of surveys (questionnaires) as a means for valuing non-market, public goods inasmuch as individuals will, when asked to value a public good, behave strategically in efforts to 'snatch some selfish benefits'; resulting biases are referred to as 'strategic bias'. This then leads to a second conclusion, viz., that in the absence of market prices reflecting (however imperfectly) individual preferences, "... we are unable to define an unambiguously 'best' state" (p. 388) in terms of a level of provision of public goods.

At about this same time, Ciriacy-Wantrup (1952) (hereafter, C-W) considered the question as to how one might obtain values for a particular class of 'extra-market' -- public -- goods, viz., public goods related to resource and environmental conservation. In this regard, C-W proposed the use of survey methods for obtaining such values:

"Individuals ... may be asked how much money they are willing to pay for successive additional quantities of a collective extra-market good ... The results correspond to a market demand schedule. For purpose of public policy, this schedule may be regarded as a marginal social revenue function." (Ciriacy-Wantrup, 1952, pp.241-42)8/

C-W considered the following five possible objections to this valuation procedure, all of which, in his view, could be reasonably overcome with the careful design of questionnaires. 9/ First, he considers the interdependence (and, therefore, non-additivity) of individual utilities, an influence which he regarded as minor and correctable by questionnaire design (C-W, 1952, p. 242). Second, he mentions the problem of 'lumpiness' in the provision of extra-market goods, a potential problem considered by him as (i) not peculiar to extra-market goods and, (ii) possibly requiring for its resolution an appeal to costs rather than benefits (C-W, 1952, p. 243) Third, he notes the potential for individuals to purposefully bias responses to interrogation. Of course, this objection is an early statement of Samuelson's 'strategic behavior' argument noted above. C-W regarded the potential bias from strategic behavior as correctable by questionnaire design and, in any case, small; of course, Samuelson regarded the issue as the "... fundamental technical difference (vis-a-vis markets.)

going to the heart of the whole problem of social economy" (Samuelson, 1954, p. 389).

The fourth objection to the use of surveys for valuing public goods considered by C-W relates to potential biases stemming from (i) the fact that other extra market goods are not considered (in a survey focused on one particular good) and (ii) the fact that the marginal utility of money is not likely to remain constant. The 'other goods' issue is considered by C-W to be of minor importance and not peculiar to extra-market goods: they "... apply also to the use of demand functions in analyzing the market" (C-W, 1952, p. 243). For 'practical' ends sought in the survey, C-W suggests that the assumption of constant marginal utility of money may frequently be realistic "... because of compensating variations in the prices of other commodities or in money income". Fifth, and finally, C-W suggests that the survey method might be regarded as too academic: the supply of extra-market goods is determined by political machinery, not by monetary valuation. Without the benefit of President Reagan's Executive Order 12291, however, C-W notes the potential contribution of value information to the decision-making process in a democratic government (p. 244).

As an aside, it is interesting to observe that the notion of 'option demand' formally introduced by Weisbrod (1964) has as its precursor C-W's observation that "... planning agents may allow for uncertainty by keeping their utilization plan flexible. This means that they may decrease the periods over which costs are sunk, avoiding obligations to pay fixed charges ..." (p. 113). Indeed, as observed by Krutilla (1967), "It must be acknowledged that with sufficient patience and perception nearly all of the arguments for preserving unique phenomena of nature can be found in the classic on conservation economics by Ciriacy-Wantrup" (p. 778).

Notwithstanding C-W's apparent optimism regarding the use of survey methods for deriving estimates for public goods values, we find no evidence of immediate efforts to develop and apply the idea. Indeed, following Samuelson's 1954 paper one finds little in the literature concerning the public goods valuation issue until the late 1960's-early 1970's. However speculative, it might seem as if Samuelson's arguments were found compelling vis-a-vis the impossibility of deriving value measures for non-market, public goods.

Three distinct lines of inquiry were introduced around the late 1960's-early 1970's which had the effect of rekindling interest in the public goods valuation issue. First, Clawson and Knetsch (1966) refined and popularized the Travel Cost Method (TCM) for valuing recreation sites. ^{10/} Second, Rosen (1974) introduced the Hedonic Price Method (HPM) as a means for valuing some classes of non-market goods. Third, the question as to the potential efficacy of surveys as a means for valuing public goods was reintroduced as a result of: (i) an experiment wherein C-W's suggestion for using surveys was implemented by Davis (1963a and 1963b) and later by Knetsch and Davis (1966); (ii) Bohm's (1971, 1972) experiments with survey methods which tested and rejected Samuelson's strategic bias hypothesis; and (iii) refinements in the survey method introduced in by Randall et al. (1974) based on the aggregate "bid curve" suggested by Bradford (1970). The structure for surveys set out by Randall et al. provides the essence of contemporary applications of survey referred to as the CVM.

The resurgence of intellectual interest in the public goods valuation

issue alluded to here is by no means attributable solely to the above-cited works. The 1960's and early 1970's were the formative years for what is now the sub-discipline of 'resource and environmental economics'. Interest in the valuation of the public good, "the environment", was stimulated by the provocative works by Krutilla (1967) and Kneese (1962), to name but two of the imaginative contributors to the air of intellectual excitement that characterized that period. Our focus on methodological lines of inquiry initiated during this period simply reflects the methodological nature of the issue of primary concern in this book.

We will not divert attention from the developments of concern regarding the CVM for a discussion of the Travel Cost and Hedonic Price Methods for valuing public goods; these methods have direct relevance for our assessments of the CVM, as is discussed below in Chapter VI. At this juncture, we wish to focus attention on developments with the CVM initiated by the works of Davis, Bohm, and Randall et al.

F. DEVELOPING THE CVM

In two ways, Randall et al.'s (1974) paper set the use of surveys, in terms of their use for estimating values for public goods, on a distinctively different track from that implied by C-W (and applied by Davis) and/or from that implied by Bohm's work. First, Randall et al. attempted to define and impose on the survey a rigorous structure designed to differentiate their use of a method whereby values were elicited from individuals (a survey) from 'ordinary' surveys. Their survey method was called a 'bidding game'. Their 'structure' was a questionnaire design wherein willingness-to-pay questions were posed within a context which draws from a market analogy: the context of a contingent market. In terms now familiar to those working with the CVM (discussed below in Section G), the 'structure' was an effort to elicit behavioral, as opposed to attitudinal, revelations of individual preferences. This structure, and its variants, are now referred to as the Contingent Valuation Method -- CVM.

Secondly, with the benefit (not afforded Davis in his earlier study) of Bohm's results which weakened Samuelson's strategic bias proposition, Randall et al. suggest the potential applications of the CVM to the task of valuing a wide range of environmental improvements -- types of public goods that extend well beyond those amenable to cross-check via other methods (e.g., the TCM with recreation demands as in the 1966 study by Knetsch and Davis) and relatively 'hard' commodities such as Bohm's Public Television commodity. In this regard, witness the 'commodity' in Randall et al.'s study: aesthetic benefits from reduced air pollution.

Randall's pursuit of these challenges was quickly joined by other scholars. Efforts to develop the promise (as it was then seen) of the CVM were focused in large part, as one might expect, on methodological problems as they related to the application of the method. In this regard, the specter of Samuelson's strategic bias proposition remained as a concern, notwithstanding Bohm's results, until appearance of Vernon Smith's (1977) report of experimental evidence that further belied the strategic bias proposition. Thus, a number of earlier CVM studies were focused on tests of the strategic bias proposition. But to test the strategic bias proposition, one needed to apply the CVM, and in efforts to apply the CVM, an ever-widening range of operational/methodological problems arose: how does one initiate the valuation process?; what is the appropriate mode of payment in which to couch the willingness-to-pay question?; what kind and how much information should be given to survey participants? 11/

As efforts to deal with operational questions of the type posed above continued, applications of the CVM were extended in innovative and imaginative ways. As examples, Daubert and Young (1981) applied the CVM for the estimation of benefits attributable to instream river flows; Walsh et al. (1978) and others applied the method to estimate option and preservation values attributable to improved water quality in Colorado's Platte River Basin; and Crocker (1984) applied the method to valuing avoided damages to forest stocks from reduced acid depositions.

Operational sorts of problems of the type mentioned above pale in significance in comparison with the problem of 'hypothetical bias', however. Regrettably, 'hypothetical bias' (HB) seemingly has many different faces -- it means different things to different people. As but a few examples, Rowe and Chestnut (1983) view HB as arising "... because respondents are predicting what their behavior would be in a hypothetical situation" (p. 408); Schulze et al. (1981, p. 158) see HB attributable to a

respondent's failure to understand all of the ramifications of a posited environmental change; Thayer (1981, p.32) seemingly views HB as potentially arising because (for unstated reasons) individuals may not behave as they indicate that they will behave (i.e. pay their WTP) in the CV interviews; Bishop and Heberlein (1979) suggest that HB may result from the fact that the CVM elicits statements of attitudes rather than intended behavior or from the fact that contingent markets are "... too artificial to provide a sufficient context for developing accurate values" (Bishop et al. 1983, p. 620); finally, although certainly not exhaustively, Burness et al. (1983) see HB as resulting from the (asserted) fact that "... the CV market precludes the derivation of values which reliably reflect the interviewee's preferences ..." (p. 675).

Obviously, from the above, the concept (or concepts) of hypothetical bias is generally intuitive and almost always poorly defined; perhaps understandably in light of the imprecision of the hypothetical bias notion, efforts by researchers to respond, via empirical tests of related hypotheses or otherwise, have been equally imprecise. 12/ An exception is found in one form of the hypothetical bias proposition which proposes that choices made under conditions where actual payments are involved will differ from choices involving hypothetical payment. This hypothesis has been stated, tested, and demonstrated as 'true' by a number of scholars. 13/ We note that this hypothesis is but one possible interpretation of the arguments of Freeman (1979a) and of Feenberg and Mills (1980) which propose that, with hypothetical payment, individuals lack incentives to incur the disutility associated with time and mental energy required to respond 'accurately' to willingness-to-pay questions. As will be argued later, however, means other than actual payment may provide incentives for accurate responses.

Given, unquestionably, that the CVM is hypothetical in character -- it involves a hypothetical market for the provision of a commodity which involves hypothetical payment -- the persistence of criticisms that CVM measures must be substantively biased is perhaps understandable; this is particularly so given the general failure by scholars working with the CVM to translate posited sources for hypothetical bias into testable hypotheses and to test them. Thus, the hypothetical bias issue, with all of its diverse, poorly defined 'faces', remains as one of the most important unresolved issues relevant for any assessment of the efficacy of the CVM as a means for estimating values for non-market environmental commodities. As we will see in the following section, the potential intuitive appeal of the hypothetical bias proposition vis-a-vis the credibility of CVM measures is reinforced by research findings in another sub-discipline.

G. RELATED RESEARCH IN OTHER DISCIPLINES

As evidenced by an examination of references in the CVM literature, scholars involved in the development of the CVM have only recently become aware of the full implications for their own work of the research ongoing in other areas of economics and in other disciplines. The attitude v. behavior issue which has long been of concern to psychologists was introduced by Bishop and Heberlein (1979). Economists' concern with mechanisms for eliciting 'true' preference revelations -- e.g., the Vickery (1961) 'second price' auction -- is only recently reflected in the CVM literature (Coursey et al. 1983), and examinations of the potential contributions to the development of the CVM from techniques derived in 'experimental economics' are at a relatively infant stage.

Also, in the area of psychology a great deal of empirical research concerning the manner in which individuals make decisions may be relevant for the CVM. As examples of the many anomalies in individual decision-making reported by Tversky and Kahneman (T-K) (1981), their observations concerning 'mental accounts' are of particular interest. T-K argue that, in making allocative decisions (regarding income), the individual may focus on groups of commodities as opposed to individual commodities. Thus, rather than allocate \$15.00 to a night at the movies, \$25.00 to an evening at the opera and \$10.00 to a day at the beach, an individual may allocate \$50.00 to something akin to an 'entertainment account'. Sub-allocative decisions are then made as the need or opportunity for recreation or entertainment arises. To the extent that individuals do think in terms of 'accounts' there may be serious implications for the CVM. In deriving a value, for example, for a specific environmental improvement (e.g., improved air quality in Denver) the obtained value may in fact apply to some more aggregate commodity (account), say environmental quality in general -- i.e., the CV measures may relate to something akin to an 'environmental account', as opposed to the specific environmental improvement serving as a 'commodity' in the CV study.

Another related line of argument that is potentially relevant for assessments of the CVM is that developed by researchers at Decision Research (Eugene, Oregon). Of particular interest is the recent work by Slovic et al. (1980). Citing recent research by T-K (see below), they argue that individuals seemingly use inferential rules, called 'heuristics', to reduce difficult mental tasks to simpler ones. Three characteristics of common heuristics used by individuals are of interest. 14/ First, individual judgements of the importance of an event, or the likelihood of its occurrence, are affected by the extent to which the event (public good) is easy to imagine or recall -- i.e., by information (in the press, T.V., etc.); this 'availability' heuristic is related to a second, 'representativeness' heuristic which will reappear below in our discussions of risk. Thus, for example, a CV study focusing on willingness-to-pay for environmental regulations on nuclear waste disposal (more generally, hazardous waste disposal) might result in seriously distorted results given recent, well-publicized events such as the Three Mile Island accident and documentaries on Love Canal. Efforts to value recreation facilities in a nearby National Park could be distorted by recent reports of crowded conditions at any recreational facility. Equally serious, values for public goods related to government actions could be distorted by exposes of official misconduct, reflecting distrust of (or distaste for) the

government in general.

Secondly, Slovic points to research suggesting that (i) individuals tend to be overconfident in their heuristics and (ii) people's beliefs, once formed, change very slowly -- judgements of 'fact' are "... extraordinarily persistent in the face of contrary evidence". (Slovic, 1980, p. 189) Thus, to the extent that individual beliefs or perceptions concerning a particular public good are fixed, the task of altering perceptions of the good -- communicating the nature of, e.g., a specific environmental improvement -- may compound the complexities involved in an individual's perception of an actual change and their valuation of that change.

Third, Slovic points to what might be referred to as a general aversion to uncertainty by individuals. Evidence from psychological research suggests that, as a means for eliminating the anxiety that attends uncertainty, uncertainty is simply denied -- a behavioral pattern vis-a-vis uncertainty noted by other authors as well. ^{15/} Results from survey methods may be seriously distorted if, indeed, individuals generally deny risk and uncertainty, particularly in studies involving public goods affecting such things as mortality and morbidity. Examples include CV studies designed to value changes in air/water quality and studies designed to value the adoption of any public policy related to health and safety.

Risk and, most prevalently, uncertainty vis-a-vis risk are common dimensions of many of the public-environmental goods of analytical interest in applications of the CVM. ^{16/} the use of the CVM to value public/environmental goods presupposes some understanding as to how individuals form values under conditions of risk and uncertainty. Underlying most analysis is the expected utility hypothesis of behavior under uncertainty combined, in a sense noted by Arrow (1982), with the implicit use of the Bayesian hypothesis wherein individuals consistently use conditional probabilities for changing beliefs on the basis of new information. A recent example of this approach is seen in a paper by Gallagher and Smith (1984) wherein, in valuing (e.g.) improved air quality in a national park, the individual perceives a 'change in air quality' as a change in the probability distribution of air quality levels to which he/she has access on any given visitor day. In the Gallagher-Smith model, "... to the extent that each individual appreciates the random nature of environmental services ..." (p.2) the individual's valuation of a posited environmental quality improvement is then based on the maximization of expected utility (within the context of state-dependent utility functions).

Another area of ongoing research of potential relevance to the CVM concerns the rationality hypothesis so basic to the bulk of economic analysis, and upon which rests the expected utility hypothesis. The rationality hypothesis has long been questioned as to its relevancy vis-a-vis empirical content and there is growing criticism as to its validity, in any operational sense, in explaining or predicting individual behavior under conditions of uncertainty. The degree of complex calculations imputed by the theory to individuals in their efforts to form valuations -- witness the weight of such calculations implied in the Gallagher-Smith application -- is belied by empirical evidence and, in the authors' minds, by intuition. As observed by Arrow,

"Hypotheses of rationality have been under attack for empirical falsity almost as long as they have been employed in economics. Thorstein Veblen long ago had some choice, sarcastic passages about the extraordinary calculating abilities imputed to the

average individual in his or her daily economic life by economists. More recently, Herbert Simon and his colleagues have produced much evidence of the difficulties of human beings in arriving at rational choices even in rather simple contexts ..." (Arrow, 1982, p.1)

Extending Arrow's reference to Simon's work, Simon notes that "When even small complications were introduced into the (decision-making) situations, wide departures of behavior from the predictions of subjective expected utility (SEU) theory soon became evident ... the conclusion seems unavoidable that SEU theory does not provide a good prediction -- not even a good approximation -- of actual behavior". (Simon, 1979, p. 506)

H. THE STRUCTURE FOR A CVM ASSESSMENT

As a result of our reflections concerning the thrusts of CVM-related research conducted over the last decade, four issues stand out in terms of encompassing questions of central importance for our efforts to assess the state of the arts of the CVM. These are: (i) questions concerning the degree to which CVM experiments have succeeded in developing questionnaire designs that mitigate or eliminate, the potential for operational-types of biases (vehicle, information, strategic biases, etc.); (ii) questions concerning the extent to which research results outside of the CVM area of research per se have been rationalized vis-a-vis their implications for the CVM -- in this regard, reference is made particularly to the areas of decision theory, experimental economics and psychology; (iii) questions concerning the pervasiveness and magnitude of biases in CVM measures. attributable to 'hypothetical bias'; and (iv) questions concerning the existence of precise standards which serve as a basis for accepting or rejecting hypotheses related to the 'accuracy' of CVM measures.

The structure for our assessment of the CVM is, therefore, one which allows sharp focus on these four sets of questions. Thus, Chapter III focuses on the questions posed in (i): CVM studies are critically reviewed with particular concern being given questionnaire design as it relates to operational biases. A review of research, and its relevance to applications of the CVM, in the area of experimental economics is provided in Chapter IV; these discussions focus on a subset of the questions implied by (ii)). The issue of hypothetical bias is addressed in Chapter V; as a part of our assessments of the many 'faces' of hypothetical bias -- the substance of question set (iii) -- we will be required to examine research results from the fields of decision theory and psychology, thereby rounding out our focus on question set (ii). Questions related to standards by which the accuracy of CV measures might be assessed (set(iv)) are, in the authors' view, of primary importance. This issue is addressed in Chapter VI. As a part of this inquiry, empirical evidence related to comparisons of CVM values with values derived from the TCM and HPM are analyzed and discussed.

Questions posed in (i)-(iv) and responses to these questions given in Chapters III - VI, will hopefully set the stage for discussions at the Assessment Conference concerning the major issue of interest in this book: the state of the arts of the CVM. As noted above, this major issue is the topic of Part II of this book.

ENDNOTES

Chapter Two

- 1) See, e.g., J. Rothenberg, 1961.
- 2) I.M.D. Little, 1952.
- 3) See, e.g., J. Rothenberg, 1961, pp.36-41. See also the conclusion in A.K. Dasgupta, and D.W. Pearce, 1978, p.89.
- 4) See Dasgupta and Pearce, 1978, p.90-93.
- 5) See W.D. Schulze, C.S. Brookshire and T. Sandler, 1981
- 6) See Dasgupta and Pearce, 1978, Chapters 2 and 4, for a discussion of this point.
- 7) See, for example, Lipsey and Lancaster (1956).
- 8) Also, "The psychological mechanism of these subjective evaluations themselves (for example, whether cardinal or ordinal differentiation of utility is involved) are neither accessible nor relevant for the observer -- that is, for objective evaluation of extra-market goods," p.85.
- 9) "Welfare Economics could be put on a more realistic foundation if a closer cooperation between economics and certain young branches of applied psychology could be established", Ciriacy-Wantrup (1952), p.244.
- 10) A letter from Harold Hotelling to the National Park Service wherein Hotelling suggests a method like the TCM is reproduced in Brown, W., A. Singh and E. Castle, 1964. See Brown, et al. (1964), for an example of competent applications of the TCM prior to Clawson and Knetsch's cited work.
- 11) For discussions of, respectively, 'starting point, vehicle and informational' biases see Schulze et al., 1981; and R.D. Rowe and L.G. Chestnut, 1983.
- 12) For example, see Burness et al. and Schulze et al., 1979.
- 13) For example, Bohm, 1972; D.L. Coursey, W.D. Schulze, and J. Hovis, 1983; P. Slovic, 1969; and Bishop and Heberlein, 1979.
- 14) Slovic et al.'s arguments focus on decisions involving risk; their arguments would seem to have broader applications however, in substance if not implied magnitudes of importance.
- 15) For example, Kahneman and Tversky (1979), and Starr, Rudman and Whipple (1976).
- 16) Given the broad class of environment 'commodities' for which option

values may be relevant, it is interesting to note that uncertainty (of purchase or use) lies at the heart of Weisbrod's definition of option value (Weisbrod, 1964). Uncertainty vis-a-vis health risks may be relevant for option value as seen in Weisbrod's example of hospitals -- a public good "... utilized infrequently by most persons and not at all by some; yet ... (providing) a valuable standby service ..." (Weisbrod, 1964, p. 474). Underlying one's option value for the hospital must be some perception of the probability -- risk -- of its use at some future date. For related discussions, see B. McNeill et al., 1981 and Lichtenstein and Slovic, 1971.

III. APPLICATIONS OF THE CVM: AN OVERVIEW OF ISSUES.

A. OVERVIEW.

In Chapter II the reader was given some flavor for the setting wherein interest in the potential of the CVM was initiated. As a part of those discussions, we noted four sets of questions that have been of primary concern for researchers involved with experimental research related to the development of the CVM. These questions were: (i) the "strategic bias" question; (ii) questions concerning the extent to which subjects in CVM experiments understand the "commodity" to be valued, as such understanding is reflected by behavior that is consistent with axioms from received theory; (iii) questions related to questionnaire design -- starting point, vehicle and information biases; (iv) questions concerning the equivalence between willingness-to-pay and willingness-to-accept values derived with the CVM; and (v), more generally, a broad range of questions concerning biases attributable to the hypothetical nature of the CVM's valuation process. In this chapter, we consider research results which are relevant for addressing questions given in (i) - (iv). Given the myriad issues relevant to an assessment of hypothetical bias and the need, in responding to related questions, for a review of research results in other disciplines, we defer to Chapter V the task of considering the hypothetical bias questions referred to in (v).

R. STRATEGIC BIAS AND THE CVM

Concern with strategic behavior on the part of economic agents can be traced historically to economists' efforts to argue for or against a mechanism or institution that would yield allocations of public goods which parallel in some sense those which would obtain in a competitive market. Wicksell (1896) suggested that:

"(if) providing the expenditure in question holds out any prospect at all of creating utility exceeding costs, it will always be theoretically possible, and approximately so in practice, to find a distribution of costs such that all parties regard the expenditure as beneficial and may therefore approve it unanimously" (Wicksell, 1896, p. 90).

Samuelson (1955) notes that Wicksell was careful to separate theoretical from practical solutions; in support of his theory of public expenditures, he argues that his theory was

"... an attempt to demonstrate how right Wicksell was to worry about the inherent political difficulties of ever getting men to reveal their tastes so as to attain the definable optimum" (p. 355)

Samuelson's categorical rejection of the possibility of obtaining "true" individual valuations of public goods due to "strategic behavior," served as a point of departure for research wherein a variety of theoretical framework and a variety of incentive-compatible auction mechanisms were developed in effects -- a la Wicksell's (1896) "approximately as in practice" dictum (p. 90) -- to resolve the problem of pricing, and thus of allocating, public goods. Authors involved in these efforts include: Groves (1973), Clarke (1971), Loehman et al. (1979), Groves and Ledyard (1977), Smith (1977, 1979), Tidemand and Tullock (1976), Bohm (1972) and Scherr and Babb (1975). In what follows, we consider the studies by Bohm (1972); Scherr and Babb (1975); and Smith (1977, 1979) wherein explicit attention is focused on the strategic behavior hypothesis.

The Bohm (1972) study involved laboratory-type experiments designed to investigate the effects on individual behavior of six alternative approaches for valuing a TV program that had not been previously shown to the public. Four of the six approaches explored by Bohm for determining aggregate willingness-to-pay required that the subject actually, as opposed to hypothetically, pay money for obtaining access to the TV program. If the aggregate stated maximum willingness-to-pay actually exceeded the cost of the TV program, the subjects were told that they would have access to the program and that they would actually pay in one of the following modes (pp. 114-15):

- (I) according to his maximum willingness-to-pay as stated,
- (II) the same fraction of the maximum stated, the fraction being equal to costs divided by the stated aggregate maximum willingness to pay,
- (III) according to one of several alternatives, the choice not yet being made,

- (IV) a given amount, the same for all individuals,
- (V) nothing.
- (VI) nothing (this was a hypothetical case).

Incentives for free riding in each of the above payment modes were viewed by Bohm as follows. For approach I, individuals will understate a willingness-to-pay -- an expectation based on Samuelson's arguments for strategic bias; for approaches II, III, and IV, Bohm argues that subjects will overstate willingness-to-pay. It should be noted that V and VI differ not only in payment modes; subjects given V and VI were also given different definitions of the "commodity" and different amounts of information.

Subjects in group VI faced a hypothetical structure quite similar to the standard CVM approach while those in group V did not. Individuals in group V "were simply asked how much they found the program to be worth at a maximum" (p. 119). Approach VI is quite similar to the contingent valuation approach as employed by Mitchell and Carson (1981) and others, which we will take up later in this chapter.

Two of Bohm's results are of interest for our discussions. First, Bohm finds that "none of these (first) five approaches ... gave an average maximum willingness-to-pay that significantly deviated from that of any other of the approaches." (Bohm, 1972, p. 112); from this, Bohm rejects the strategic bias hypothesis. Second, Bohm finds that the sixth approach did produce a hypothetical willingness-to-pay significantly above average valuations obtained in the other five approaches. Such differences lead Bohm to conclude that:

"... when no payments and/or formal decisions (emphasis added to distinguish group VI from where payments were also not required) are involved ... this ... may be seen as still another reason to doubt the usefulness of responses to hypothetical questions, in general, and of ordinary polls (emphasis added) to guide political decision making with respect to public goods in particular." (p. 125)

We should note that the weight of Bohm's results, at least as regards his conclusions concerning the effects of hypothetical payment, may be diminished somewhat by results reported by Mitchell and Carson (1981). Mitchell and Carson contest Bohm's conclusion in this regard for two reasons. First of all, after deleting an unusually large bid, the authors found the group VI mean bid to drop substantially, to the point where the statistical difference between groups III and VI vanished. Secondly, the authors found that income in group VI was higher, than in group III, leading to the possibility of an income effect explaining the differences found by Bohm between the group VI and other group bids.

Scherr and Babb (1975) examined the theoretical pricing system constructs proposed by Clarke (1971) and Loehman, et al. (1979), in a controlled experimental setting for the pricing of two public goods: a concert and a library fund. Scherr and Babb's rationale for testing the Clarke multi-part pricing system and the Loehman-Whinston average incremental cost pricing system was the assertion that:

"If the predictions of the theory deviate from the observed behavior in this setting, one may begin to question the possible linkage of the theory to real world behavior." (p. 36)

Scherr and Babb's focus on strategic or free-rider behavior is a bit curious in the following ways. The hypotheses tested by Scherr and Babb were stated in terms of whether the Clarke and Loehman-Whinston systems would inhibit free-rider behavior and, implicitly (it would appear; see pp. 45-48), they assume that "the subjects could have been free-riding under the voluntary systems" (p. 46). The authors conclude that neither of the "... proposed pricing systems (neither the Clark nor the Loehman et al. pricing systems) inhibited free-rider behavior of the subjects" (p. 47). However, as mentioned above, this analysis was predicated on the assumption of free-riding in the voluntary system. Thus, if the voluntary system did not lead to free riding by the subjects, then the result that: "There were not significant differences in the demand levels associated with the pricing system" (p. 47) would appear to cloud our attempts to determine whether Scherr and Babb "found" or even "inhibited" free-riding in the experiments utilizing the alternate pricing schemes. This confusion is seen in their assertion that:

"The outright offer was the simplest of all situations in that the subjects only had to indicate what part of the 50 cent allotment they wish to donate to sponsor four concerts (books). The opportunity to be a free-rider could not be clearer than in this situation. Yet the outright offers were significantly higher than comparable offers under even the voluntary system, about 45 percent higher." (Scherr and Babb, 1975, p. 45)

The authors noted that the proposed "voluntary system closely corresponds to commonly experienced methods of contributing to community projects ..."
(Scherr and Babb, 1975, P. 46) Further,

"The proposed pricing systems may not have inhibited free-rider behavior because there was not a great deal of such behavior to inhibit. The debriefing suggested that few subjects attempted to free-ride." (p. 46)

The authors add:

"A different population might contain a larger proportion of people who would attempt to be free-riders and thus improve the chances that the proposed pricing systems would inhibit such behavior." (p. 46)

This last statement is especially interesting in that it suggests only a fraction of a population might free-ride; thus to observe this fraction the sample population must be increased. The experimental arguments set out by Scherr and Babb do not suggest pervasive strategic behavior by individuals.

We next briefly consider results from two studies by V. Smith (1977, 1979) which address the strategic bias hypothesis. Smith (1977) reports results obtained in laboratory experiments wherein incentive-compatible auction mechanisms are used in eliciting subject's valuations of public goods. Smith (1979) reports results from a series of experiments utilizing the

Groves-Ledyard (G-L) incentive-compatible tax rule for valuing public goods:
On the basis of these studies, Smith concludes that:

"What emerged from this paper, ..., is that practical decentralized processes exist for the provision of public goods. Some of these processes lead to optimal or approximately optimal allocations. If there are a few such processes there must be thousands -- some better, some worse, some cheaper, some clearer." (Smith, 1979, p. 62)

"Why do they not (individuals in the experiments) exhibit the more 'sophisticated', 'strategic' behavior postulated by Hurwicz and Ledyard-Roberts? I think it is because there are significant direct (and indirect) opportunity costs of thinking, calculating, and signaling which makes strategizing uneconomical." (Smith, 1977, p. 1136)

Thus, results from Smith's laboratory experiments belie the notion that individuals behave strategically in response to public good valuation questions.

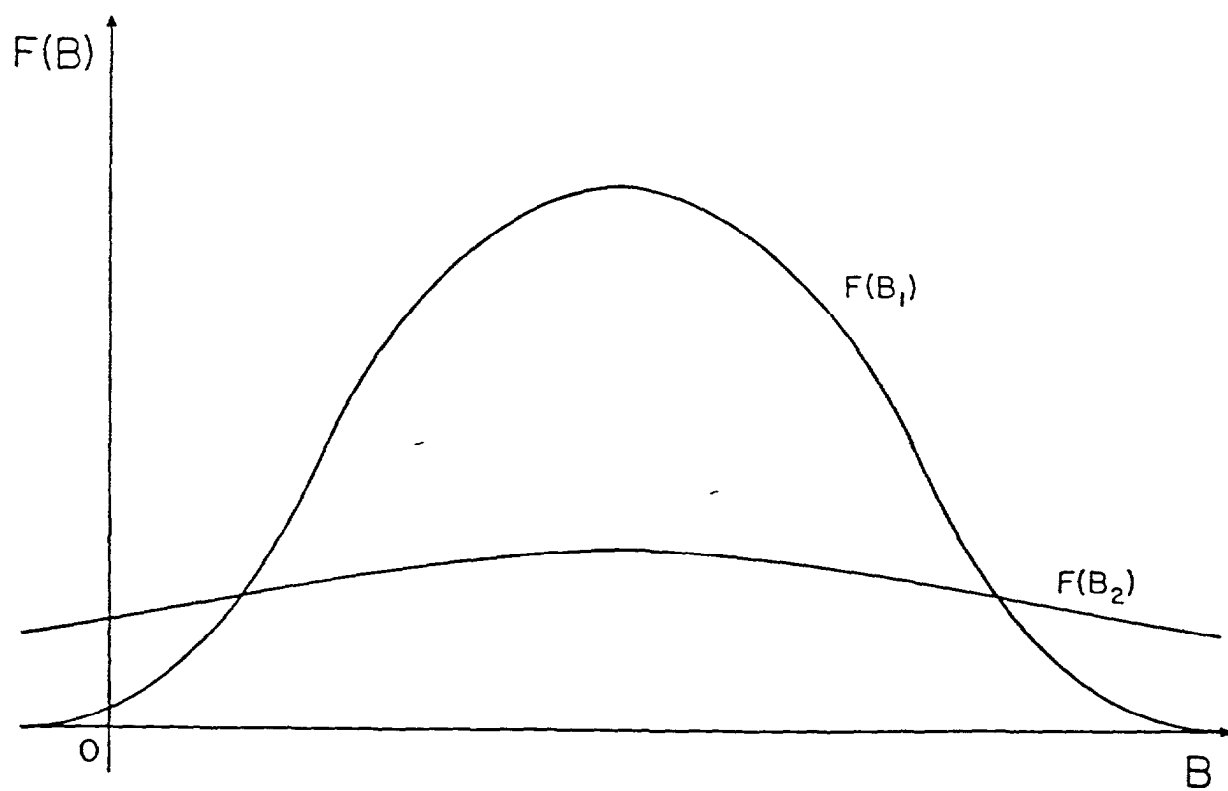
The studies cited above involve controlled laboratory experiments which focus on the strategic bias question. This question has also been addressed in CVM studies. Results from three of those studies are of particular interest for our discussions -- the studies by Brookshire, Ives, and Schulze (1976); Rowe, d'Arge, and Brookshire (1980) and Mitchell and Carson (1981).

Brookshire et al.'s (1976) study was based upon the following arguments. Consider the individual whose 'true' bid is different (either higher or lower) from other subjects. In order to behave strategically, a substantially large false bid (relative to the sample mean bid), that deviates from the individual's "honest" bid, would have to be given in order to affect the overall sample mean bid if the strategically-behaving individual is to effectively impose his/her preferences on other subjects. For an "environmentalist", when environmental preservation is at issue, infinity may be the upper bound on his/her bid, while for a "developer" the relevant bid may be zero. Thus, given the assumption that "true" bids are distributed normally, as illustrated by F(B.1) in Figure 3.1, the Brookshire et al. "test" of strategic bias involves the inspection of the actual bid distribution. That is, the greater the occurrence of strategic bidding, the flatter the distribution of bids, as illustrated by F(B) in Figure 1. Thus, if CVM bids included a large number of zero and high bids, thereby producing a "flat" distribution of bids, strategic behavior is assumed to be indicated.

Based upon the argument that bids are distributed normally and that strategic behavior will serve to flatten the distribution, results from the authors' application of the CVM lead them to conclude that "the results of the survey ..., do not lead to the conclusion that strategic behavior was prevalent among the recreators interviewed at Lake Powell" (Brookshire et al., 1976, p. 340).

Rowe et al. (1980) approached the problem of testing for strategic bias differently. Their study involved willingness-to-pay and willingness-to-accept measures for preserving alternative levels of air quality in the Four Corners Region of the Southwest. Subjects from whom CVM valuations were obtained were also asked questions related to their attitudes about environmental issues. Subjects were then classified as: conservationist,

Figure 3.1: Examples of Bid Distributions



semi-conservationist, middle-of-the-road, semi-developer, or developer. Significant correlation between bids and additional dummy variables was interpreted as being indicative of the presence of strategic bias. The authors concluded that:

"..., the results yielded no significant interactive dummy variables, hence no strategic bias for either the ES or CS bids."

The authors further conclude

"... that if zero and very large bids are closely analyzed and possibly rejected, strategic bias, if it exists, has a negligible effect upon the bid distribution." (Rowe et al., 1980, p. 15)

Using a bid distribution argument analogous to that used by Brookshire et al. (1976), Mitchell and Carson (1981) investigated the prevalence of strategic bias in CVM bids for improvements in national water quality. Mitchell and Carson's approach differed from that of Brookshire et al., however, in the following way. Mitchell and Carson use average U.S. income distribution (rather than Brookshire et al.'s "normal distribution") as a "normal" distribution in analyses concerning the flatness of the distribution of bids from a CVM. Their analyses result in the following conclusions:

"The overall shape of the (bid) distribution is not flat. It approximates a log normal distribution, a distribution similar to that reported by Brookshire, et al. (1976) in their Lake Powell study, and to the distribution of income in the United States. Since income is a strong predictor of people's willingness to pay for water quality, as we will see in Chapter 5, we conclude that the distribution does not suggest strategic bias." (Mitchell and Carson, 1981, pp. 4-10)

"Eighty-three percent of those who gave amounts greater than zero fall into our 'normal' category. Those in the extreme categories are divided, with 10 percent giving 'high' amounts and 7 percent willing to pay low amounts. We conclude that those at the extremes are relatively few in number and rather evenly balanced." (Mitchell and Carson, 1981, pp. 4-13)

Thus, Mitchell and Carson do not find evidence of strategic bias in the results of their application of the CVM.

Results from experimental laboratory and CVM studies concerning efforts to test the strategic bias hypothesis reviewed above do not support the hypothesis. Of course, these results cannot be interpreted as definitive evidence that subjects will not behave strategically in applications of the CVM. As noted earlier, one may criticize structures for questions and information used by Bohm in his experiments. Scherr and Babb's conclusions may be weakened by their basic assumption of free-riding behavior in voluntary exchange systems. The weight of Smith's findings may be challenged by an appeal to the simplified artificial setting of laboratory experiments (an issue discussed below in Chapters IV and XIII). Rowe et al.'s conclusions are not supported by a compelling argument as to why correlation between environmental attitudes and bids would indicate

strategic bias -- if strongly conservationist attitudes carry through to budget-related preferences, lack of significant correlation between attitudinal variables and bids might be indicative of strategic or other biases (as opposed to their contrary interpretation). Finally, Brookshire, et al. and Mitchell and Carson's studies, which look to "flat" bid distributions as manifestations of strategic bias, may leave some unconvinced as to: why "biases" might not be more or less normally distributed across surveyed populations and/or be sufficiently biased bids incomes so as to result in a distribution of strategically biased bids that approximates the distribution of strategically biased bids which might be directed at studies which have focused on the strategic bias issue notwithstanding, the authors find impressive the consistent lack of success in identifying such biases in these studies. Thus, while acknowledging the absence of a basis in these studies. Thus, while regard, we suggest that at a minimum, a basis does exist for diminishing the "priority" position in research agendas that the strategic bias hypothesis has enjoyed for the past decade.

C. AXIOMATIC BEHAVIOR AND CVM VALUATIONS

The economists' theory of value assumes that individuals have perfect knowledge over all states of the world, alternative actions and post-action states of the world. In homier terms, the individual is aware of all possible goods/services (and their prices) that he/she might buy, as well as savings alternatives, his/her income and his/her preferences regarding all combinations of purchased goods/services and savings. Based on such perfect knowledge, the individual selects purchases such that equimarginal conditions obtain; i.e., the ratios of marginal utilities to prices for all purchased commodities are equated.

If subjects interviewed in applications of the CVM behave -- in terms of their formation of willingness-to-pay responses -- as individuals are presumed to behave in market environments, the above-cited axiom from utility theory might be used as a basis for deriving testable hypotheses concerning the extent to which the CVM does, as assumed, "simulate" the market environment. Several authors have taken this tack, testing one or more of the following hypotheses. In what follows, define V as a subjects' stated willingness-to-pay in a CVM study. Let $V(y)$, $V(g)$ and $V(b)$ be values obtained under conditions where the subject is asked to reveal his/her income and monthly expenditure patterns as well as to identify the expenditure category which must be reduced if the subject is to actually pay his/her stated bid for the CVM commodity ($V(y)$); the subject is "reminded" of "other goods" which he/she might purchase in lieu of the CVM commodity ($V(g)$); and where a repetitive bidding process is used -- "would you pay \$1 more?" ($V(b)$). The following hypotheses are considered.

(a) $V = V(y)$: i.e., bids obtained wherein the individual's "budget constraint" is made explicit, are the same as bids obtained without explicit mention of the budget constraint. Equality in (a) is taken to imply that subjects in CVM experiments do, as required by the theory of value, consider income and other goods trade-offs in formulating willingness-to-pay responses.

(b) $V = V(g)$: i.e., bids obtained with and without "reminding" subjects of expenditure alternatives are the same. Equality in (b) is taken to imply that subjects, in valuing the CVM commodity, are cognizant of all states of the world as assumed in value theory.

(c) $V = V(b)$: i.e., the bidding process does not affect bids. Equality in (c) is taken to imply that a subject's initial bid is a preference-research, maximum willingness-to-pay for the CVM commodity.

Studies wherein hypothesis (a) was tested include those by Schulze et al. (1983), Sorg and Brookshire (1984), Blumberg (1984) and Walbert (1984). For all experiments included in these four studies, the authors fail to reject the hypothesis $V = V(y)$. Thus, the authors of those works conclude that CVM values are indeed formulated within a mental context in which subjects are aware of income trade-offs implied by their stated willingness-to-pay.

Hypothesis (b) is tested in three experiments reported in Schulze et al. (1983) as well as in Blumberg (1984) and Walbert. Generally, the authors' results imply the rejection of the hypothesis $V = V(g)$, i.e., the explicit introduction of other alternative goods (typically other public goods) does result in a significant change in the subject's willingness-to-pay for the CVM commodity. Curiously, the authors seemingly view this result as "good news" as well as bad news (see Schulze et al. (1983), Chapter 1). The good news is that, with the introduction of other goods,

the reduction in "expenditures" on the CVM commodity (reflecting, one must suppose, the allocation of expenditures to one or more of the "other" goods) is consistent with the axioms of utility theory. The bad news is that the perfect information assumption is seemingly violated; one must then wonder what, the effects on CVM valuations might be of explicit mention of still other alternative goods/services that the subject may not have considered in the CVM valuation process.

Finally, Schulze et al. (1983), Blumberg (1984), Walbert (1984) and Desvousages et al. (1984) report experiments which include tests of (c). It is generally the case that $V = V(g)$ is rejected -- the bidding process results in significantly higher bids for the CVM commodity. This result, particularly in Schulze et al. (1983), is interpreted as categorically implying the critical role of the bidding process in inducing preference research on the part of CVM subjects which is required for a subject's formulation of a maximum willingness-to-pay for the CVM commodity.

Results from the above-described tests are obviously somewhat mixed vis-a-vis demonstrations that the CVM valuation process approximates "real," market-like behavior. Thus, the comfort that one might take from demonstrations that budget constraints are seemingly operative in a CVM subject's formulation of an offered willingness-to-pay may be dissipated by demonstrations that such subjects are not cognizant of other, possibly competitive, public goods -- this issue concerning the range of information considered ("processed") by individuals in forming values, will be pursued at greater length in Chapter V. In terms of the necessity of including a bidding process in CVM applications, the evidence in this regard appears compelling to the authors. As will be shown, results from experimental work in other areas, especially in experimental economics (Chapter IV) support the argument that repetitive bidding-like trials are required in the CVM as a means for assisting the subject to learn the valuation process and in inducing preference research.

D. BIAS ISSUES RELATED TO THE DESIGN OF CVM QUESTIONNAIRES

Three of potential biases in CVM value measures which may be attributable to the manner in which CVM questionnaire are designed have been dominant in terms of eliciting concern by researchers involved with the development of the CVM. These bias issues, discussed below, are typically described by the rubrics: starting point bias, vehicle bias and information bias.

1. Starting Point Bias.

Randall et al. (1974) suggested that respondents be asked to respond "yes" or "no" to a question of the form: Would you continue to use this recreation area if the cost to you was to increase by X dollars?" (p. 135). By varying the amount \$X given to different groups of subjects, a demand curve for the recreation area could then be derived. A problem arose, however, concerning the rationale for choosing any value(s) for X and the potential that such choices would result in biased responses (i.e., "starting point" bias). Two possible sources for starting point bias have been identified. First, the starting bid may suggest (incorrectly) to the individual the approximate range of "appropriate" bids or costs for providing the environmental good. Thus, the individual may respond differently depending on the magnitude of the starting bid. Second, if the subject values time highly, boredom or irritation may set in with any lengthy iterative bidding process. In consequence, if the suggested starting bid is substantially different from actual willingness-to-pay, the subject may be unwilling to go through a lengthy process of searching preferences required for arriving at a maximum willingness-to-pay. It was hypothesized that the effect of these two types of starting point bias would substantially influence the accuracy of contingent valuation measures and, therefore, the usefulness of the approach for the assessment of preferences.

Several studies have explored whether starting point bias exists by examining the effects of alternative starting points (Randall, Grunewald, et al., 1978; Brookshire D'Arge Schulze and Thayer, 1981; Brookshire, Randall, and Stoll, 1980; Rowe, d'Arge, and Brookshire, 1980). Other studies have explored the effectiveness of alternative valuation mechanisms in avoiding a starting point bias -- an example is the payment card, on which a wide range of dollar values is listed. In the case of the payment card, the choice of a starting bid is left up to the subject in that the subject chooses his/her "starting point" from the values given on the payment card. Rowe et al. (1980) utilized starting bids of \$1, \$5 or \$10, and introduced these values as an independent variable in the estimation of a bid equation as a statistical test for starting point bias. The coefficient was significant and positive, indicating that choice of a starting bid significantly influenced mean bids. Rowe et al. conclude that "the effect of increasing the starting bid was approximately \$0.60/month on a \$1.00 increase within the \$1.00 to \$10.00 range examined" (p. 12). In passing, we note the limited range (\$1.00 to \$10.00) of starting points used by Rowe et al., a characteristic of their study which has led later writers to question the strength of their conclusions.

Brookshire et al. (1980), in a study of wildlife values, employed starting points of \$25, \$75, and \$200. Brookshire et al. fail to find a significant relationship between starting points and final bids: "the

hypothesis that final value data were influenced by the initial bids posited to respondents ... (is) rejected at the .05 level of significance." (p. 64)

Brookshire, d'Arge, Schulze, and Thayer (1981) explored starting point bias in a contingent valuation study of air quality in Los Angeles. Subjects in twelve communities in the Los Angeles area were surveyed in an attempt to determine willingness-to-pay for improvement in air quality. Three starting points -- \$1, \$10, and \$50 -- were used in the questionnaires. This resulted in three potential comparisons of starting point effects on mean bids: (1) \$1 to \$10; (2) \$1 to \$50; and (3) \$10 to \$50. The authors tested the null hypothesis of equality across bids from each starting format, ignoring all other potential effects on bids. The null hypothesis of equality was rejected. Thus the authors found no evidence of starting point biases and concluded that such biases may not be a major problem for applications of the CVM.

Thayer (1981) conducted a contingent valuation experiment wherein starting points of \$1 and \$10 were used. Three different tests for starting point bias were undertaken: 1) a comparison of mean bids from differing groups of subjects; 2) estimation of a linear bid equation

$$\text{Final Bid} = a - B(s)$$

where $B(s)$ is the starting point; and 3) estimation of a generalized bid equation inclusive of social and economic variables. Thayer's results were as follows. The mean bid comparison indicated "no difference between average bids differentiated by starting point even at the 10 percent significance level" (Thayer, 1981, p. 41). The estimated linear equation showed "the coefficient on starting point ... approximately equal to -0.02, implying that a one dollar increase in the starting bid will cause a two-cent decrease in the bid, an insignificant effect as indicated by the negligible t-statistics" (Thayer, 1981, p. 41). Finally, utilizing the generalized regression (which included social and economic variables), "the most noteworthy feature of the equation is that the coefficient on the starting point was not significantly different from zero" (Thayer, 1981, p. 42).

While the above-cited studies suggest that starting point biases may be of minimal importance for applications of the CVM, results from a number of other studies suggest otherwise. Thus, significant effects on mean bids from starting bids -- i.e. starting point bias -- are reported in research conducted by, e.g., Mitchell and Carson (1984) and Boyle *et al.* (1984). (The authors acknowledge Mitchell and Carson's suggestions in this regard; see Appendix to Chapter XIII below.)

As noted above, concern over the problem of starting points also led researchers to consider alternative mechanisms for eliciting initial bids, most notably, the use of a payment card. Experiments with payment cards included, in many cases, the use of iterative bidding processes discussed above in sub-section C. The implied rationale for tying iterative bidding to payment cards was seemingly the notion that a subject's initial choice from a payment card may not reflect the subject's maximum willingness-to-pay; thus, iterative bidding is assumed to provide incentives for the subject to search his/her preferences for the maximum amount he/she would pay for the CVM commodity.

Sorg and Brookshire (1984) and Schulze *et al.* investigated the relationship of payment card bids and bids obtained with iterative bidding. Mean bids and standard errors from those studies are presented in Table 3.1. Examination of Table 3.1 indicates that the iterative bidding approach yields measures up to 40 percent higher than initial bids taken from the payment card. As noted above, the authors interpret these results as suggesting that iterating initial bids is an important element in the contingent valuation methodology.

Table 3.1

Iterative Bidding and the Payment Card Approach.

Average Bid
(Standard Errors)

Using:

Commodity:	Iterative Bidding	Payment Card	Sample Size
Visibility at the Grand Canyon (a)	\$9.20 (11.54)	\$5.69 (7.21)	64
National Water Quality (a)	\$8.71 (11.11)	\$5.50 (8.4)	56
Containment of Hazardous Waste (a)	\$25.35 (36.43)	\$16.02 (20.78)	163
Elk Wildlife Encounter (b)	\$55.50 (36.43)	\$44.50 (20.78)	20

(a) See Schulze et al. (1983) for further details.

(b) See Sorg and Brookshire (1984). Their bids are for the situation where the hunter typically sees 10 elk per day.

Finally, two studies consider interactions between the interviewer and the subject as a possible explanation of the wedge between payment card values and the iterative bidding values noted in Table 3.1. Sorg and Brookshire (1984) found no statistical difference between mean bids obtained via payment card (no iteration) in a personal interview format and mean bids obtained via payment card in a mail questionnaire. Schulze, Brookshire et al. (1983) reach a similar conclusion in a study of ozone effects in Los Angeles. CVM values for reduced ozone concentrations were obtained from in-person interviews (no iterative bidding) and mail responses. Referring to tests of the hypothesis that interview bids equal mail survey bids, the authors conclude that:

"In no case can this hypothesis be rejected at the .05 level, and even at the .10 level the hypothesis can be rejected only in Orange County." (Schulze, Brookshire et al., 1983, p. 5.41)

Thus results from research to date do not provide a basis for unequivocal conclusions concerning the relevance of starting point bias in CVM studies. Furthermore, we have noted that the use of the payment card format without iterative bidding yields significantly lower values than those derived with an iterative format. Thus, available evidence suggests the desirability Using iterative bidding procedures in CVM applications wherein payment cards are used.

The role of iterative bidding procedures in CVM applications is further developed below in Chapters IV and VI.

2. Vehicle Bias

When willingness-to-pay questions are posed to subjects in an application of the CVM, the questions are typically posed within a context that describes how the subject would pay his/her offered payment; as examples, payment via tax payments, entrance fees (to recreation areas), utility bills, or simply higher prices for goods and services. Considerable attention by CVM researchers has been given to potential biases in willingness-to-pay measures that are associated with the choice of a mode of payment or "payment vehicle." For example, if a subject has an aversion to higher taxes, the subject might understate his/her willingness-to-pay for an environmental commodity if such payment must be made through higher taxes. Resulting biases are described as "vehicle biases." Essentially, one finds two possible sources or manifestations of vehicle bias discussed in the literature. First, it is argued that vehicle bias is demonstrated when either mean bids or the recorded number of protest votes varies significantly with the choice of vehicle. Secondly, drawing from economic theory wherein substitution possibilities differ with alternative payment mechanisms, when a payment vehicle allows the individual to substitute over a wider range of current commodity purchases, it is argued that the bid for any given CVM commodity should be higher.

Vehicle bias has been examined by a wide variety of researchers including Randall et al. (1978); Brookshire, Randall and Stoll (1980); Rowe et al. (1980); Brookshire, d'Arge, Schulze, and Thayer (1981); Greenley et al. (1981); Loehman et al. (1981); Cronin (1982); and Daubert and Young (1982). In the wildlife study by Brookshire, Randall and Stoll (1980), the authors utilized hunting license fees and utility bills as bidding vehicles, and tested the null hypothesis that bids were unaffected by the choice of payment vehicle. The results were not conclusive, as is illustrated by the following:

"The hypothesis that final bids ... were influenced by the choice of bidding vehicle (a component of the bidding scenario) was rejected at the 0.1 level of significance. Nevertheless, it was observed that refusal to bid, with WTP formats, occurred in six of fifty cases with a 'utility bill' vehicle, but in none of fifty-eight cases which used a 'hunting license fee' vehicle. Negative comments in the 'feedback' section occurred more frequently with the 'utility bill' vehicle".
(p. 484)

Rowe et al. (1980) utilized utility bills and payroll deductions as payment vehicles. The payment vehicle was treated as an independent dummy variable in an overall bid regression where a bid based upon a utility bill was designated 0 while a payroll deduction bid was designated 1. For equivalent surplus bids, the coefficient on the dummy variable was positive and significant (i.e., the t-statistic was 3.05). For compensating surplus bids, the coefficient on the dummy (payment vehicle) variable was negative and not significant (i.e., the t-value was -.696). Thus, their results were inconclusive as to the existence of vehicle bias.

Brookshire, d'Arge, Schulze, and Thayer (1981), in an air quality study in Los Angeles, conducted a test of means between bids with a monthly utility bill vehicle and a lump sum payment vehicle. The authors report the

following conclusion.

"the null hypothesis of equality of the mean total bids irrespective of the bidding vehicle cannot be rejected for Montebello, Canoga Park, Encino, Huntington Beach, Newport Beach, Pacific Palisades, Palos Verdes, and Redondo Beach. However, for Irvine, Culver City, La Canada, and El Monte, we reject the null hypothesis, at least at the 90% confidence level, for the total bid. The principal reason for these differences seems to stem from the aesthetic bids." (Brookshire et al., 1981, p. 148).

Greenley, Walsh and Young (1981), in a recreation study of the South Platte River Basin in Colorado, utilized a general sales tax and a residential water sewer fee as bidding vehicles. From tests as to the influence of payment vehicles on bids, the authors suggest:

"... that willingness to pay for water quality was quite sensitive to the method of hypothetical payment. Residents sampled reported willingness to pay only about one-fourth as much in water-sewer fees as in sales tax for the option value of water quality. Respondents were more reluctant to participate in the water-sewer bill estimation procedure and may have perceived inequities. Everyone including tourists, pays sales taxes; whereas only property owners and indirectly renters, pay water-sewer bills. Moreover, recent experience with escalating water-sewer fees may have resulted in understatement of willingness to pay for water quality" (p. 671).

Finally Daubert and Young (1982) conducted a study focusing on recreation demand for maintaining instream flows on the Cache la Poudre River in northern Colorado. The two payment vehicles used in the study were: increments in county sales tax on consumption expenditures; and entrance fees for three recreation activities (fishing, shoreline, recreationists, white water kayakers). Front tests for vehicle bias, the authors state that "The estimated bid functions for the three recreation activities were statistically different for each repayment obligation; sales tax marginal benefits always exceeded entrance fee values" (p. 672).

Thus, we find rather persistent evidence that supports the vehicle bias proposition -- the choice of a payment vehicle would seem to be an important determinant of values derived with the CVM. What is not apparent from the received literature is how one might go about eliminating such biases -- how one identifies a "neutral" or unbiased vehicle. Questions related to this issue will be addressed by participants at the Assessment Conference, described below in Part II of this book.

B. Information Bias.

Information bias is one of the more difficult sources of bias to define with any degree of precision; different researchers have used and explored different notions of such biases. The broadest definition was suggested by Rowe et al. (1980) as "A potential set of biases induced by the test instrument, interviewee, or process, and their effects on the individual's responses". In principle, the different aspects fall into three categories. First, those biases, such as starting point or vehicle bias, which have been discussed earlier. Second, the order in which information is collected or

elicited from the respondent is hypothesized to affect the mean bid -- a potential bias described by other as a "sequencing bias" (see Brookshire et al., 1981). Third, information bias is argued to result from the quality and quantity of information given to subjects in the CVM.

Rowe et al. (1980) examined the third view of information bias described above via giving groups of subjects information which differed in quality. Following a subject's bid, the subject was given (randomly chosen) mean bids from other subjects, after which the subject was allowed to alter his/her initial bid. All the subjects were told that they would pay the overall mean bid. This second element allowed the respondents to revise their bid based upon "new" information (average bids by others) if they desired to do so. Thus, the reader sees in this aspect of Rowe's test for information bias the form of a test for strategic bias. Rowe et al.'s test for information bias involved the construction of a dummy variable where a value of 0 was assigned if the subject was not told the mean of other's bids, and 1 if such information was provided. The test result shows the coefficient to be negative and significant (the relevant t-statistic was -4.54). The authors concluded that:

"The effect of prior information concerning previous mean bids, which were stated to have been in the \$1.00 to \$1.50 range, was equally significant This result suggests that if the individual is given sufficient information and their true bid exceeds the stated mean bid, they illustrate a form of the classical free-rider behavior by bidding less than their maximum willingness to pay. However, note that the formal structure of the iterative bidding technique need not provide the necessary information to create this incentive" (Rowe et al., 1980, pp. 12, 14).

Brookshire et al. (1981) obtained bids for the elimination of aesthetic and health (acute and chronic) effects related to air quality. Subjects were asked to value alternative combinations of reduced (i) aesthetic, (ii) acute health and (iii) chronic health effects. Their analyses focused on the impact on bids for a particular effect of the sequence in which the effects were introduced. The two alternative sequences used were: 1) aesthetic, aesthetic plus acute, and aesthetic plus acute plus chronic or 2) acute, acute plus chronic, and acute plus chronic plus aesthetic. This allowed for the examination of two hypotheses. First, individuals will bid differently for reduced aesthetics (or acute health effects) depending upon where in the sequential bidding process the aesthetic (or acute) effects are introduced. Second, sequence (1) will result in a cumulative bid (for the reduction of all effects) that differs from sequence (2). The cumulative, or total, bid for all effects assumes additivity with respect to the subject's preference structure related to air quality effects. The authors found that effect-specific bids, as well as total bids, obtained with sequence (1) were significantly different from those obtained with sequence (2). Thus, they conclude that information bias as it relates to the sequence in which information is presented to subjects may be of real concern to those involved with the development of the CVM.

Cronin (1982), in a water quality study conducted along the beaches of the Potomac River designed a survey to examine the effects of different quantities of information on subjects' willingness-to-pay. A subset of subjects was informed that "it will help you to know that the average

household in the D.C. Metropolitan area is paying about \$30 per year to maintain the existing water quality ..." (p. 5.4). All other subjects were not given this information. Cronin concludes:

"While it is difficult a priori to hypothesize the directional bias that additional information might induce on elicited bids, . . . comparisons involving the information-no-information situation all indicate substantial differences between respondents provided with cost estimates and those not provided with such estimates". (Cronin, 1982, p. 6.11)

As an aside, Cronin also informed one group of subjects that their bid would affect local taxes while others were told that the federal government would bear the costs:

"... respondents informed that their bid will impact their local taxes express a willingness to pay significantly lower than do respondents informed that the federal government will bear the costs" (Cronin, 1982, pp. 6.100).

Related to our discussions of strategic bias above in sub-section B, Cronin argues that these results are indicative of strategic behavior.

A similar test was conducted by Schulze et al. (1983) in their "Policy Bid Experiment". The authors attempt to discover whether factual information on the current level of expenditures for environmental regulations would affect the initial bid given by subjects for a "new" regulation to control hazardous wastes. Prior to posing willingness-to-pay questions, one half of the sample was informed of the approximate amount they were currently paying in higher taxes and prices for the current state of environmental quality; the other half was not given this information. The authors report a failure to reject the hypothesis of equality between the bids of the two groups -- evidence of information bias was not found. They conclude:

"It would appear that, in offering contingent values for our policy commodity, individuals may be, in general terms, cognizant of the existing state of environmental regulations and the cost of maintaining this state." Schulze et al. 1983, p. VI-49)

E. WILLINGNESS TO PAY VS. WILLINGNESS TO ACCEPT

Received theory establishes the argument that the amount of money that individuals are willing to pay (WTP) for marginal increases in consumption states available to them should approximately equal the amount of money that they are willing to accept (WTA) for an identical decrement in such consumption states. This argument is developed by Willig (1976) for price changes and by Randall and Stoll (1980) and Takayama (1982) for quantity changes. As a part of these theoretical arguments, income effects, typically viewed as "small" are shown to drive a "small" wedge between measures of WTP and WTA for a given individual.

In contrast with theoretical axioms which predict small differences between WTP and WTA, results from CVM applications wherein such measures are derived almost always demonstrate large differences between average WTP and WTA. Results from fifteen CVM experiments by eight groups of researchers are given in Table 3.2. As seen in Table 3.2, derived measures of WTA are consistently larger -- on the order of three to five times larger -- than measures of WTP.

To date, researchers have been unable to explain in any definitive way the persistently observed differences between WTA and WTP measures. Appeal is made to assertion of possible cognitive dissonance (Coursey et al., 1983) on the part of subjects, or to possible effects arising from voluntary exchange (WTP) as opposed to involuntary exchange (WTA) structures, but we know of no studies wherein posited causes of WTA-WTP differences have been systematically examined. WTP and WTA measures shown in Table 3.2 are typically elicited from different groups of subjects -- rather than from one subject -- but income differences between groups of subjects are generally not sufficiently large to warrant the attribution of WTA-WTP differences to an income effect. Thus, at this point in time all that can be said is first, we have observed differences -- large differences -- between WTA and WTP measures obtained in applications of the CVM; and secondly, we have little more than intuitive conjectures as to why such differences persist in CVM results. Setting aside such anomalies found in results from CVM applications, some insight as to a rationale for WTA-WTP differences may be gained from ongoing research in experimental economics. An overview of such research is given below in Chapter IV; we thus defer further discussion of this issue to Chapter IV's review of experimental economics.

Table 3.2

Measures of WTP and WTA a

Study		WTP	WTA
Hammack and Brown (1974)	(1)	\$247.00	\$1044.00
Branford, Knetsch and Mauser (1977)	(2)	48.00	120.00
		22.00	93.00
Sinclair (1976)		35.00	100.00
Bishop and Heberlein (1979) (b)		21.00	101.00
Brookshire, Randall and Stoll (1980)	(1)	43.64	68.52
	(2)	54.07	142.60
	(3)	32.00	207.07
Rowe, d'Arge and Brookshire (1980)	(1)	4.75	24.47
	(2)	6.54	71.44
	(3)	3.53	46.63
	(4)	6.85	113.68
Hovis, Coursey and Shulze (1983)	(1)	2.50	9.50
	(2)	2.75	4.50
Knetsch and Sinden (1983)	(1)	1.28	5.18

a All figures are in year-of-study dollars. The bracketed numbers refer to either the number of valuations received or the number of trials (in experiments) conducted.

b Carson and Mitchell (1984) reestimated Bishop and Heberlein's results with contrary conclusions.

F. CONCLUDING REMARKS

Experimental efforts to develop the CVM as a tool for deriving estimated values associated with public/environmental goods have enjoyed substantial progress in in many areas. Improvements have been made in some areas of questionnaire design -- e.g., in the use of visual aids for communicating to subjects the substance of hypothetical changes in the environment (see Schulze et al., 1983) -- and in the development of imaginative applications of the method to a wide variety of environmental commodities (e.g., Walsh et al., 1978). Also, as noted above in sub-section B, experimental research with the CVM (and research in other fields) has provided an empirical perspective regarding "strategic bias" in CVM results wherein the potential for such biases is no longer a source of preoccupation for CVM researchers -- strategic behavior by subjects in applications of the CVM is no longer considered inevitable nor is the potential for related bias thought to be a matter for primary concern.

Less progress has been made in term of responding to other questions related to the efficacy of the CVM for its intended uses. While CVM subjects seemingly consider income constraints in their formulation of valuation responses, their valuation of a given CVM commodity may be substantively affected by: "reminders" of other, substitute, public goods, which they might wish to "purchase"; alternative modes of payment (payment vehicles); and different (quantitatively and/or qualitatively) sets of information concerning the CVM commodity. When payment cards are used in lieu of starting points, existing evidence points to the necessity of using an iterative bidding process as a part of the CVM application if measures of a subject's maximum willingness to pay for a commodity are to be obtained. Finally, large differences between WTA and WTP measures derived from applications of the CVM persist and remain unexplained.

While CVM research specifically directed at questions of the sort described above has not produced definitive results, it would be premature at this point in our discussions to suggest state of the arts conclusions as to the implications of research results reviewed in this Chapter. Insights relevant to assessing the issues discussed in this Chapter are found in results from research in other disciplines and in results from CVM research which is directed at the broader question as to the nature of "hypothetical bias" in values derived with the CVM. These topics are addressed in the following three chapters. Thus, a formulation of our tentative (pre-Conference) conclusions regarding the implications of research reviewed in this Chapter for the state of the arts of the CVM must await discussions in Chapter VI where results from our more comprehensive review of multidisciplinary research are used in efforts to suggest state of the arts conclusions.

IV. EXPERIMENTAL ECONOMICS: IMPLICATIONS FOR THE CVM.

A. INTRODUCTION

As noted in Chapter III, the contingent valuation approach has been used to generate willingness-to-pay functions for a large and diverse set of consumer goods. The principal concern remains that answers to hypothetical survey questions concerning value may be biased -- they may not reveal individual preferences in any meaningful way. As originally expressed by Bohm (1972), the fact that respondents do not actually pay for the provision of the public good in question gives rise to problems in interpreting reported values. As argued above, while not necessarily having an incentive to exhibit free-rider behavior, subjects may simply have no incentive to "tell the truth" and may easily be influenced by spurious, irrelevant factors such as a desire to please the surveyor or the desire to avoid socially unacceptable responses.

Researchers have attempted to reduce the potential for these irrelevant factors in CVM applications by making survey questions as realistic as possible. This has led Davis (1963) and Randall et al. (1974) to construct so called bidding game surveys wherein the valuation process is initiated with the subject's response to an initial Starting bid after which the interviewer begins a process of asking for increasingly higher commitments for payment until the respondent indicates that he or she would not pay more for the public good than the last price quoted by the interviewer; when "high" initial values are used, and initially rejected by the subject, the initial value is incrementally lowered until the subject indicates a willingness-to-pay.

Another approach, described in detail in Chapter III, which has been used by Mitchell (1981) and Schulze and Brookshire et al. (1983) in the valuation process, involves the use of the payment card. In this type of survey, the subject is asked to circle that amount of money from a set of alternatives printed on the payment card which most closely represents his or her maximum willingness-to-pay. Schulze et al. (1983a) used the results of three public goods studies to show that willingness-to-pay obtained from the iterative bidding approach significantly exceeds willingness-to-pay obtained from the payment card approach. For the studies given in Table 3.1 the iterative bidding approach yields value measures that are about 40 percent higher than those obtained with the payment card approach. Why would or should we expect these differences? Which is the appropriate technique to employ?

Randall et al. initially used an iterative bidding approach because they hypothesized that such a process might be more "market-like" to subjects and could, therefore, simulate a competitive auction experience. In fact, auction results from laboratory experiments have shown that even when it is theoretically in the immediate best interest of an individual subject to reveal his/her maximum willingness-to-pay, the auction process yields values which reflect full willingness-to-pay only after a series of iterative learning periods (Cox, Roberson, Smith, 1982). 1/ This would suggest a priori that an iterative bidding survey scheme might be expected to outperform the payment card approach.

A second unresolved problem in the contingent valuation approach is the unexpectedly large value difference obtained for both private and public goods in willingness-to-pay (WTP) and in willingness-to-accept (WTA)

compensation studies. Theoretically, questionnaires designed to ask an individual for payment to acquire a good should provide similar results as questionnaires designed to ask an individual how much compensation is required to give up the same good.² However, results from the studies compiled in Table 3.2 of the previous chapter serve to document the large differences between WTP and WTA measures obtained in CVM studies. The questions then arise: should one use WTA or should one use WTP measures of value in contingent valuation studies? which, if either, corresponds most closely to values which are "true" in the sense of meaningful revelations of preferences? In what follows, we consider results from experimental economics as they provide insights regarding these important questions.

B. METHODOLOGICAL DEVELOPMENTS IN EXPERIMENTAL ECONOMICS.

Contingent valuation surveys are designed to collect field data relevant for social policy analysis using alternative survey instruments (questionnaires). Each of the instruments has its own set of rules and therefore causes a specific set of individual messages about the public good whose level of provision is to be increased or decreased. The survey method exercises control over changes in the institutional rules for allocating a public good, but it offers little or no control over the incentives which may affect the subjects' valuation of the good. A researcher may propose a new questionnaire design and test that design in the field. However, lacking control or information concerning preferences, the results of that survey cannot be unambiguously interpreted. Evaluation of each survey's results is complicated by the classic problem of underidentification. Field experiments must be interpreted in terms of prior assumptions regarding individual preferences and behavior as they are implied by the rules of the survey. However, the fundamental objective behind a laboratory experiment in economics is to create a manageable "microeconomic environment in the laboratory where adequate control can be mandated and accurate measurement of relevant variables guaranteed" (Wilde, 1980, p. 138). As noted by Smith (1977), control and measurement can only be measured in relative terms, but undoubtedly are much more precise in the laboratory than in the field.

The most important concept in the evaluation of an allocative system, and the concept which has driven institutional theorists, is that of "incentive compatibility". An institution's rules are incentive-compatible "... if the information and incentive conditions that it provides agents are compatible with the attainment of socially preferred outcomes This means that the rules specified in the institution in conjunction with the maximizing behavior of agents yields a choice of messages which constitutes an equilibrium whose outcomes are (socially desirable)." (Smith, 1982, p. 927).

Vickrey (1961) published the first article in which a mechanism for achieving optimal allocations in laboratory settings was proposed. His sealed-bid auction mechanism had the property that each participant had a dominant bidding strategy to truthfully reveal demand. Vickrey's fundamental and path-breaking result has recently enjoyed a renaissance and has precipitated considerable attention on the design of demand-revealing mechanisms: Shubik (1975); Dubey and Shubik (1980); Cox, Roberson and Smith (1982); Forsythe and Isaac (1982); and Milgrom and Weber (1982).

Most of this literature analyzes a model in which a single indivisible object is to be sold to one of a group of potential buyers. Each bidder has preferences defined over the object and over risk but not necessarily over the value to other bidders. The auction is assumed to be a noncooperative game played by the bidders.

Two kinds of auction mechanisms have been considered in the theoretical literature, oral auctions and sealed-bid auctions. In oral auctions an exchange of messages occurs between individuals according to a set of rules of negotiation. A contract can then occur. In an English auction, bids are announced by the buyers, a bid remains standing until a new higher bid replaces it, and the auction stops when an auctioneer decides that no higher bid will be forthcoming from the buyers. In a Dutch auction, price is set initially "high" and then lowered automatically in

increments until a price is accepted by one of the buyers; the acceptance terminates the auction. In sealed-bid auctions, individuals submit messages to a seller or a representative of the seller who then determines outcomes based upon a set of pre-announced rules. In a first price auction the buyer who submits the highest bid receives the object and must pay his bid. In a second price auction the highest bidder also receives the object but only pays what the second highest bidder bid. 3/ Several interesting results emerge from the theoretical consideration of these auctions. 4/

- 1) In first-price auctions the optimal individual bid is less than the value of the auctioned item. That is, an individual has no incentive to reveal demand.
- 2) The first-price auction does not imply Pareto optimal allocations.
- 3) Conclusions concerning the first-price auction also apply to Dutch auctions.
- 4) In second-price auctions the optimal individual bid is equal to the value of the auctioned item. That is, an individual's incentive is to reveal demand.
- 5) The second-price auction implies Pareto optimal allocations.
- 6) Conclusions concerning the second-price auction also apply to English auctions.

Based upon the results of 12 experiments conducted by Coppinger, Smith and Titus (1980) and 780 experiments conducted by Cox, Roberson and Smith (1982), 5/ the above implications were supported for groups of size four or greater except that first-price and Dutch auctions did not appear to be exactly isomorphic. The deviant results for groups of size less than four were conjectured to be due to a failure in the assumption of noncooperation. An important conclusion from these studies was that not all subjects in a second-price sealed-bid auction realize that their dominant strategy was to offer bids equal to their maximum willingness-to-pay; some subjects never realize this. Others require a period of time over a sequence of bidding games to "learn" the strategy. Coppinger, Smith and Titus "... question whether any meaningful one-shot observations can (therefore) be made on processes characterized by a dominant strategy equilibrium" (1980, p. 21). It appears that the desirable properties of second-price auctions -- elicitation of "true" preference revelations -- can be obtained, but sometimes only in a limited sense, after the subject has had time to experience the operation of the valuation mechanism.

Why does the second-price auction have such nice theoretical properties and the first-price auction not have them? Vickrey (1976) has posited the following intuitive explanation:

"The essence of these cases that admit of the achievement of a Pareto-optimal result seems to be the extent that the participants have a choice as to participating or not, it is an all-or-nothing choice. There can be no strategic holding back (of demand): for an individual to

hold back is to achieve a zero gain for himself." (Vickrey, 1976, p. 15)

This general result has led researchers to consider the properties of more complex multiple unit auctions. Engelbrecht-Wiggans (1980) has shown that, when more than one unit is auctioned in a single sealed-bid auction, the desirable properties of demand revelation are not achieved. Individuals will tend to understate willingness-to-pay. If each person can only bid on one unit however, the desirable properties of the second-price auction will result (Vickrey, 1976). The performance of auction mechanisms which include more complex bidding, such as a sealed-bid auction involving a single price for a multiple number of units or a sealed-bid auction in which the individual submits a different bid for each unit, is examined by Dubey and Shubik (1980); Palfrey (1980); Coursey and Smith (1982); and Miller and Plott (1983).

The implications of these results from private good auction theory for the design of contingent valuation surveys are as follows. First, they provide insights concerning how true valuations might be elicited. Individuals must be placed in an "all or nothing" situation in the questionnaire where no strategic holding back can help them. If the questionnaire can be designed in such a manner that a single unit or a single unit per individual is to be hypothetically auctioned off in a second-price fashion, then more demand-revealing behavior, and therefore information about true valuations, should be expected to occur. Secondly, an iterative auction framework is suggested. Because of the "learning period" required for incentive-compatible demand revelations found in experiments with the second-price auction, individuals also should be placed in a survey situation which provides them with tentative information about allocation before results are finalized. 6/

The question as to just how the auction mechanisms developed in experimental economics might be applied to public goods valuations in the CVM setting, warrants specific attention. In a series of papers, Smith (1977, 1979a, 1979b, 1980); Ferejohn, Forsythe and Noll (1979a, 1979b); and Ferejohn, Forsythe, Noll and Palfrey (1982) have considered the application of auction mechanisms to the problem of valuing public goods. 7/ Such applications involve the design of a process initially suggested by Groves and Ledyard (1977). In a public good auction individuals submit desired quantities of the commodity and the cost share or contribution for the commodity that they would voluntarily accept. Each individual is told the average group quantity and his or her share of total cost given the contributions of others in the group. Each individual then has the right to veto or agree to the tentative results. Group agreement prevails if and only if each individual agrees upon the outcome and the group covers the cost of the proposed amount of the public good. If agreement is reached, then each individual receives the public good and must pay his or her cost share.

The veto condition means that we have a tatonnement process in the sense that no contracts can occur until all individuals in the group are in equilibrium or agreement. This provides at least a partial solution to the problem of free-riding or the incentive to contribute less than true maximum willingness-to-pay. One individual can veto the results of the auction even if every other individual in the group agrees about a given quantity and distribution of cost shares.

A number of experimental and field applications of auction mechanisms similar to those described above have been conducted. Experimental applications include those by Smith (1979a, 1979b, 1980); Ferejohn (1982); and Coursey and Smith (1982); field applications include those by Bohm (1972); Ferejohn and Noll (1976); and Scherr and Babb (1975).

Results from these studies also suggest how an iterative auction framework can be integrated into a questionnaire framework. An iterative or sequential survey can be combined with a tatonnement voting process. Such a unanimity requirement is used in the London gold bullion market (Jarecki, 1976) and has been found to improve efficiency in private as well as collective allocation mechanisms (Smith, Williams, Bratton and Vannoni, 1982; Smith, 1982; Coursey and Smith, 1982; and Miller and Plott, 1983).

C. RECENT APPLICATIONS OF LABORATORY METHODS RELATED TO CVM DEVELOPMENTS

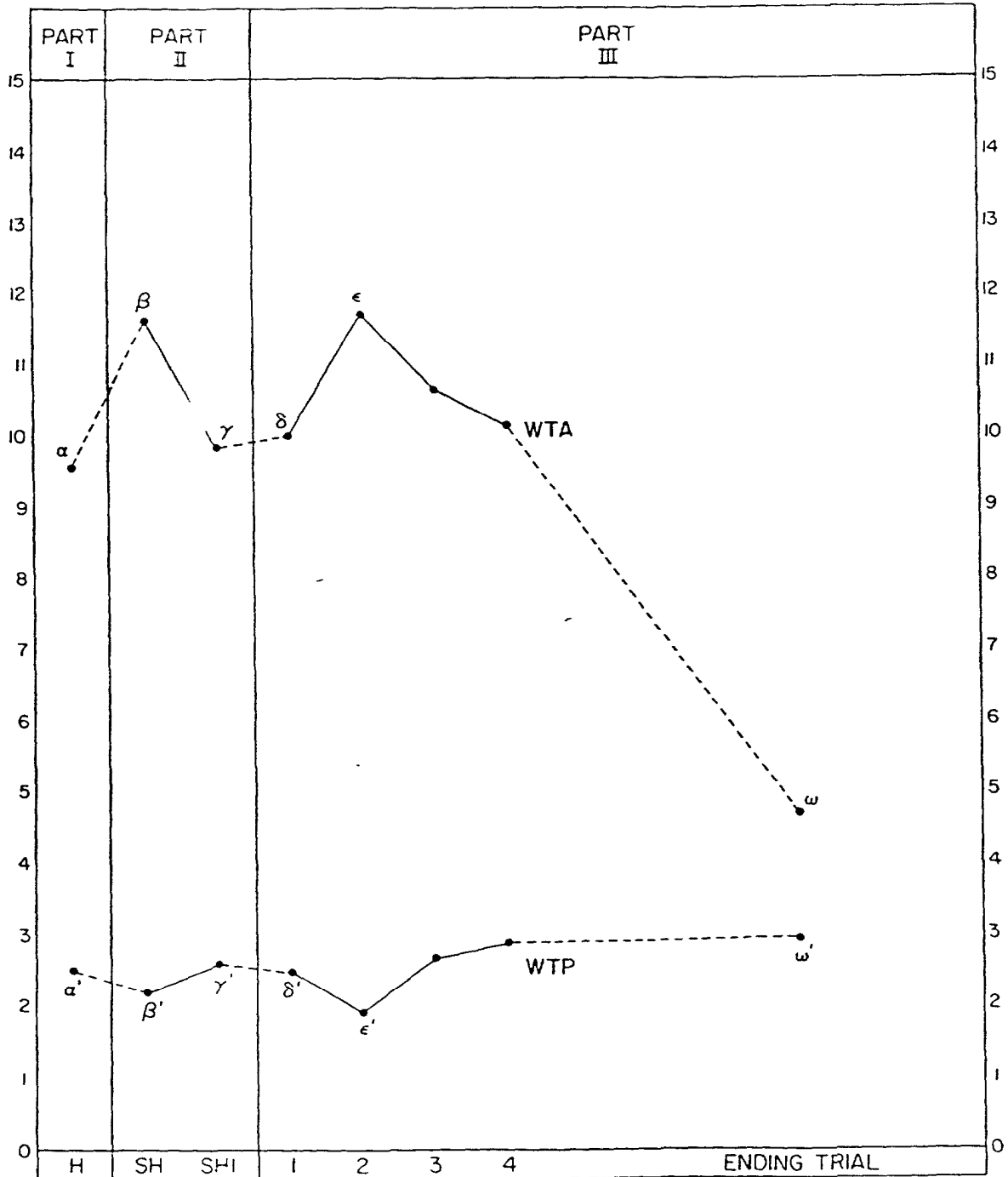
Two recent experiments were motivated at least in part by assessment-related questions in the CVM literature -- primarily to WTP-WTA differences discussed above in Chapter III. The first experiment, conducted by Knetsch and Sinden (1984), demonstrated that the large disparity between willingness-to-accept (WTA) and willingness-to-pay (WTP) measures of value is found to exist in cases where actual (as opposed to hypothetical) payments are made in the laboratory. Unfortunately, the Knetsch and Sinden experiment did not use a demand-revealing mechanism such as the Vickrey second-price auction described above. They argue that the large disparity between WTA and WTP measures of value may be due to what psychologists term "cognitive dissonance."

The second experiment, conducted by Coursey, Schulze and Hovis (1983), addressed several questions of concern for CVM developments: issues concerning the large disparity shown to exist between WTA and WTP measures of value and issues concerning the efficacy of payment cards and the iterative bidding process as methods for eliciting hypothetical payments. Given the potential importance of these issues for our later discussions, the Coursey, Schulze and Hovis (hereafter, CSH) experiment is described in some detail as follows. Individuals were assumed to have a state-dependent utility function which included income and also exposure to an unpleasant (bitter) taste experience. The experiment was designed to determine how individuals value this unusual experience from both the perspective of accepting payment to endure the experience and from the perspective of paying to avoid a bitter-tasting experience. The bitter substance used in the experiment, sucrose octa-acetate (SOA), has long been used by psychologists in taste experiments and provides a carefully controlled, safe, but unpleasant experience (Green, 1942 and Linegard, 1943).

The CSH experiment consisted of three parts. In Part I, each subject was asked to provide either a hypothetical WTA or a WTP for tasting SOA based on a verbal description of the substance. In Part II, subjects were allowed to sample a few drops of SOA and were again asked for either WTA or WTP. Respondents were then allowed to change their earlier (Part I) bid and an iterative bidding procedure was used to determine maximum WTP (or minimum WTA). In Part III, groups of eight, who were originally asked the WTA questions, participated in a Vickery auction for a fixed supply of four one ounce cups of the SOA. Low bidders were then actually compensated to taste the substance. For groups originally asked the WTP questions, a similar Vickery auction was held for not tasting the substance and high bidders actually paid their offered amounts to avoid tasting SOA. Presumably, the well documented demand-revealing properties associated with the competitive Vickery auction should have provided "true" values in the form of individual bids.

The results of the CSH experiment are summarized in Figure 4.1. First, note that as one moves from left to right across Figure 4.1, WTA and WTP move in opposite directions through each and every phase of the experiment. Hypothetical WTA and WTP values (given as average values across individuals) are initially far apart (points α and α' , respectively). This result is consistent with the existing literature on field applications of the survey approach for valuing public goods (Bishop and Heberlein, 1979 and Rowe

Figure 4.1: Overall Average Experimental Responses



Each point represents overall average of the thirty-two individuals who participated in each of the WTA and WTP experiments.

et al., 1980. Surprisingly, actual experience with the commodity (tasting SOA) in Part II drives hypothetical WTA and WTP values further apart (points β and β'). The iterative bidding process results in WTA and WTP values which converge (points γ and γ'); obviously this suggests that the iterative procedure may be of some value. As the Vickrey auction begins in Part III (points δ and δ'), opening bids for WTA and WTP are similar to, but further apart than, the iterated hypothetical bids. In the second auction trial (ϵ and ϵ') WTA and WTP diverge, possibly due to efforts by some subjects to employ dynamic trial strategies not addressed in the static Vickrey models. In early trials individuals may not initially understand that the best strategy is to reveal true values but, ultimately, WTA and WTP values do indeed converge (points ω and ω'). This convergence is, however, strongly asymmetrical in that the WTA measure of value "collapses" downward under the competitive market-like experience of the auction while WTP trial values show only modest upward movement.

Final auction measures of WTA (point ω) and WTP (point ω') are statistically similar. However, although hypothetical WTA (e.g., the point γ) is not statistically similar to WTA obtained in the auction (point ω), hypothetical willingness to pay (point γ') is statistically similar to WTP obtained from the auction (point ω').

Results from the CSH experiment suggest the following conclusions. First, the lack of significant differences between WTA and WTP measures in this experiment may be attributable to the demand-revealing nature of, and learning experiences in, the Vickrey auction. This result is consistent with economic theory and suggests that the observed divergences between hypothetical measures of WTA and WTP may result mainly from lack of a market-like environment.

Second, hypothetical WTA measures of value are likely to be biased upwards vis-a-vis what we would interpret as true values obtained from a market-like auction. Psychological factors may of course explain this bias. However, economists might argue that opening WTA bids might well be biased upwards for simple strategic bidding reasons.

Third, hypothetical WTP measures of value may correspond more closely to true (final Vickrey auction) value than do WTA measures.

D. VALUATIONS UNDER UNCERTAINTY CONDITIONS: RELEVANT RESULTS
FROM LABORATORY EXPERIMENTS

The experimental economics literature provides insights to still another set of issues of relevance for our assessment of the CVM viz, issues concerning individual behavior under conditions of uncertainty. In this regard, Grether and Plott (1979) have documented the phenomenon of "preference reversal" for the case in which individuals face a choice between two lotteries. Consider the following example: Lottery A has a high probability of a low monetary reward. Lottery B has a lower probability of a higher monetary reward. Grether and Plott demonstrate convincingly that the same individual will often choose Lottery A over Lottery B but assign a higher monetary value to B than to A. Preferences, as determined by the pattern of choice, are reversed when expressed in monetary terms.

Grether and Plott did not use repetitive trials wherein, as in the CSH experiment, subjects might "learn" dominant strategies. Thus, Pommerhehne, Schneider and Zweifel (1982) argue that since the Grether and Plott study was a "one-shot" experiment and since "judging gambles is cognitively difficult" (p. 570), then in a second trial of an experiment structured similarly to Grether and Plott's, the frequency of preference reversals would be reduced. This in fact did not occur in their experiment to test this hypothesis. As an aside we note that two trials may still have been insufficient for subjects to have "learned" dominant strategies--in the above described experiment by CSH, four non-binding learning trials and up to ten total trials were allowed. In another related experiment conducted by Reilly (1982), it was shown that additional information, including a detailed explanation of expected values and monetary incentives, reduced the frequency of preference reversals. However, such reversals still occurred frequently.

The preference reversal issue relates to the larger question concerning the efficacy of the economists' expected utility (EU) model in describing individual behavior under conditions of uncertainty. Results from research conducted by psychologists (reviewed below in Chapter V) seriously challenge the "rationality" precepts underlying the EU model -- a challenge which finds support in the research of decision theorists (Arrow, 1982; Simon, 1979) and experimental economists. However, one finds in the experimental economics literature reported results which suggest that predictions from the expected utility model may be satisfied asymptotically after many experimental trials with subjects. Plott and Sunder (1982), in an experiment examining the rational expectations model, found that:

"There seems to be no doubt that variables endogenous to the operation of these markets served to convey accurately the state of nature to otherwise uninformed agents. We can conclude that rational expectations models (based on maximization of expected utility) must be taken seriously as not universally misleading about the nature of human capabilities and markets." (p. 692)

The implications of this result for CVM may be that when individuals are dealing with a new, highly uncertain, commodity; the survey instrument

may not be able to supply enough of a learning experience, in a reasonably short time frame, to allow an asymptotic approach to rational expected utility-maximizing behavior.

These experimental results effectively support the psychologists' arguments that serious problems may exist for traditional economic value theory where a high degree of uncertainty is present.⁸ Although some progress is being made in developing an alternative model of value under uncertainty (see for example, Chew and MacGrimmon, 1979), however, it is premature at this date to adopt a new economic-theoretical perspective.

E. AN EXAMPLE: REVELATION OF COMPENSATING INCOME VARIATION.

In order to illustrate some of the points made in the previous sections we consider the problem of constructing two different survey instruments which attempt to reveal how much individuals are willing to accept in order to have a factory move into their physical environment. The first survey proposed is structured more or less along the lines of current contingent valuation practice. The second is structured along the lines of current experimental economics practice, using a hypothetical Vickrey second-price auction.

Suppose that the environment consists of $i = 1, 2, \dots, I$ individual economic agents who have utility functions defined over income, Y_i , and Q_i , a "bad" commodity such as the smoke produced by the factory. Thus,

$$U_i = U_i(Q_i, Y_i)$$

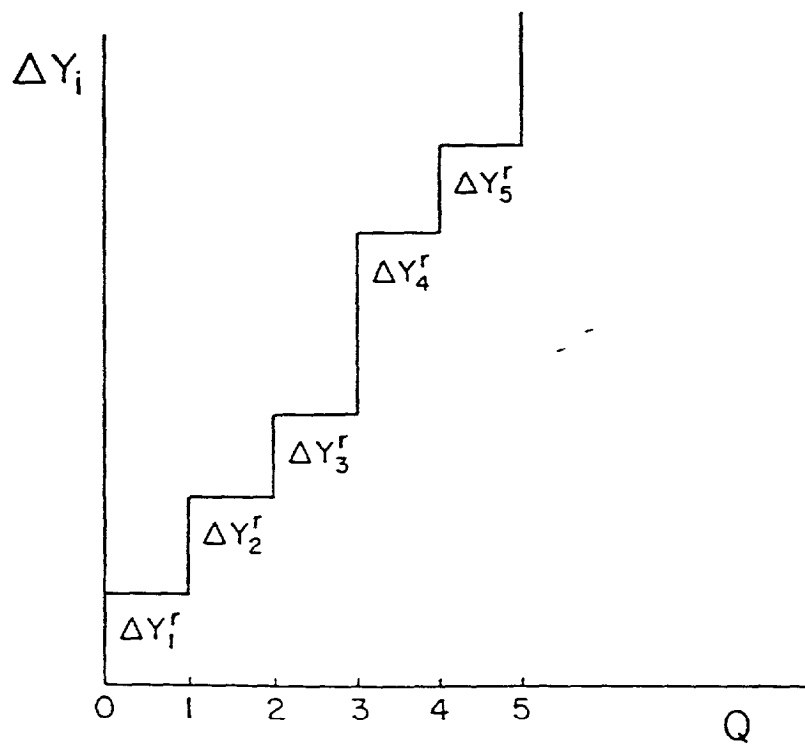
is individual i 's utility function with $U_i/Y_i \geq 0$ and $U_i/Q_i \leq 0$ for all i . Suppose that there exists an income compensation ΔY_i which would just make an individual i indifferent to a choice between a smoky environment and extra income and a clean environment with no extra income. Or, ΔY_i is implicitly defined by $U_i(Y_i + \Delta Y_i, 1) = U_i(Y_i, 0)$. Thus, ΔY_i is i 's willingness to accept monetary payment for the smoke produced by a nearby factory.

Suppose now that the ΔY_i are rank-ordered from $i = 1, 2, \dots, I$, and that $\Delta Y_1 < \Delta Y_2 < \dots < \Delta Y_I$. Then this ranking defines a compensating income variation supply function 9/ (See Figure 4.2). This curve may also be thought of as the supply function for pollutable locations. Assume for simplicity that the factory produces an integer $N < I$ total units of pollution and that the maximum consumption of Q is one unit per individual. Each individual who is affected by the factory consumes one unit of pollutant and each individual who is not affected by the factory consumes zero units of the pollutant. The situation described can be imagined as a cloud of smoke which, as it grows in size (N), envelops more and more homeowners (individuals) who surround the factory which emits the smoke. The problem facing the economist is to conduct a survey to determine the damages done by a given factory which produces N units of smoke. In what follows, we consider two institutional approaches for estimating such damages.

1. Solicited Compensating Variations.

The first approach in response to this problem might involve the construction of a survey which solicits or asks each i to submit a message m_i which is his or her willingness to accept an income compensation offer (ΔY_i) for one unit of Q ; i.e., $m_i = \Delta Y_i$. This would require only one period of data collection and analysis. Allocation of one unit is made to the N individuals who submit the lowest willingness-to-accept offers. For these individuals, $U_i = U_i(Y_i - m_i, 1)$. All other individuals j receive no units of Q , and for this group $U_j = U_j(Y_j, 0)$. The problem with this institution is that a dominant strategy involves the individuals' asking for an infinite income compensation. 10/ There is no incentive for an individual to provide the surveyor with any accurate information concerning his/her actual willingness-to-accept-payment except perhaps a desire to be honest, which may conflict with any auction-like experience the respondent may have had. This theoretical result is consistent with the large difference between willingness-to-accept and willingness-to-pay previously

Figure 4.2: Group Willingness to Pay Function
($I = 5$ assumed)



shown in Table 3.2.

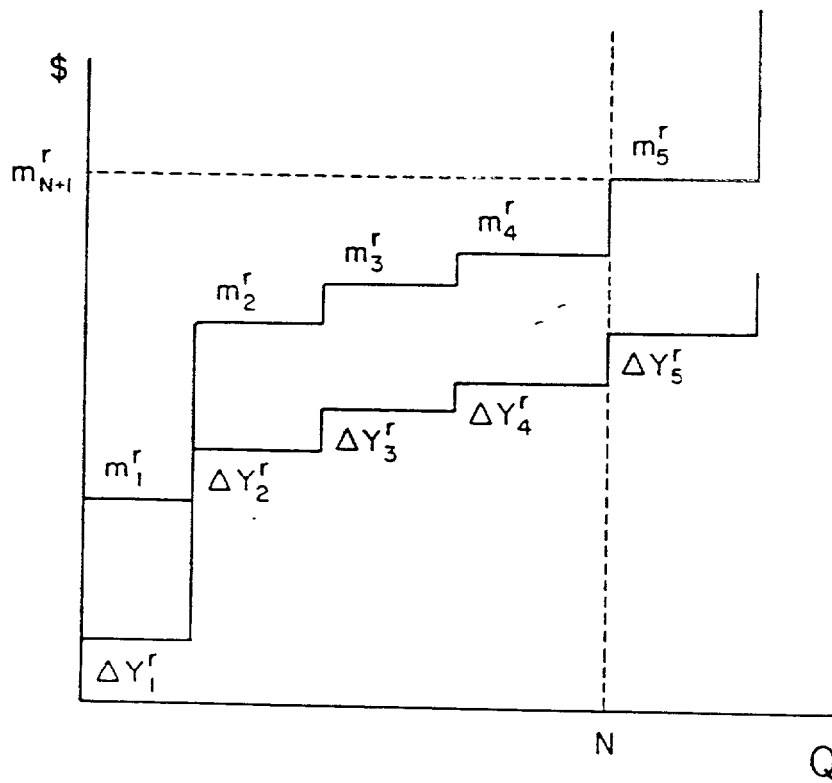
2. Tatonnement Version of the Second-Price Auction.

Now consider an alternative iterative survey. During each trial t ; $t = 1, 2, \dots, T$, let each individual i submit a message m_i which is his or her willingness-to-accept an income compensation offer for one unit of Q . Tentative allocation would then occur according to the following rules: First, the offers m_i would be ranked from lowest to highest such that $m_1 \leq m_2 \leq \dots \leq m_T$. A reigning offer price for all accepted offers m^* would be determined according to rules of second-price auction. Thus, $m^* = m_{N+1}$ (see Figure 4.3). For this first trial round, if $m_i < m^*$ then an individual would be compensated with a payment of m^* and would have to consume one unit of pollutant; for this group it would be true that $U_i = U_i(Y_i + m^*, 1)$. If $m_i \geq m^*$ then an individual would receive no compensation and would consume zero units of the pollutant; for this group $U_i = U_i(Y_i, 0)$.

These results from the first trial of the survey would then be put to a vote. All members of the group who were allocated one unit of the pollutant would vote on whether to finalize the allocation results for that trial. If all voted "yes" then everyone would realize their allocations. If at least one individual voted "no", thereby vetoing the results of the trial, then a new trial would be conducted. A second survey would be administered. The survey and voting processes would continue until a unanimous agreement occurred or until a maximum number (T) of trials had been conducted. In that case, some terminal (perhaps random) allocation procedure might be invoked.

Notice that this survey instrument incorporates three elements which theoretically and empirically should allow it to outperform the first survey. It is a second-price auction, iterative learning effects are permitted to occur, and it includes a tatonnement process. Its primary disadvantage over the simple survey lies in the cost of performing multiple trials. The two surveys might easily be compared in the laboratory. Monetary values can be induced which reflect the compensating income required for each individual to hypothetically consume a fictitious pollutant. In addition, more complicated allocation mechanisms can be constructed and tested for cases where individuals may consume more than one unit of the pollutant or where the pollutant is a pure public good or externality. Similarly, the performance of the relatively simple hypothetical iterative bidding game and other intermediate mechanisms can be contrasted to the Vickrey second-price auction. Value measures derived from each institution can be assessed for accuracy through laboratory experiments.

Figure 4.3: Vickrey Auction of N Units
($N=4$, $I=5$ assumed)



F. CONCLUSIONS.

We have argued in this chapter that a dynamic, iterative survey mechanism may well need to be employed in the design of CVM survey instruments in order to improve the accuracy of responses. Furthermore, due to the current inaccuracy of hedonic and travel cost approaches for valuing public goods, the least cost method, in our view, for testing alternative survey instruments is to use laboratory experiments. The objective of these experiments should be the development of the most simple survey design which gives accurate responses in terms of eliciting preference revelations from subjects. Several questions are implied by the discussions in this chapter: is a complex iterative voting procedure required; how fast will such a procedure converge to "true" values; what is the effect on incentives of relaxing the unanimity voting feature for large groups; can a contingent valuation mechanism be constructed which overcomes cognitive difficulties observed when individuals face an uncertain situation for the first time? All of these operational questions can at least qualitatively be answered in an experimental laboratory setting.

ENDNOTES

Chapter Four

- 1) The example cited refers to a second-price Vickrey sealed-bid auction. It is a dominant strategy equilibrium for each individual in such an auction to bid full value or reveal demand for the single unit sold in each period. At best, it usually takes subjects a few periods to realize this. Some individuals never totally reveal demand. See Cox, Roberson, Smith (1982) for details,
- 2) The difference between the two measures in theory is due to an income effect. This income effect is argued to be "small" in most cases. See Willig (1976).
- 3) These descriptions are meant to be brief. For a detailed description of the four basic auction types see Cassady (1967) or Coppinger, Smith and Titus (1980).
- 4) All are derived in Cox, Roberson and Smith (1982). See also Milgrom and Weber (1982).
- 5) See also Smith (1967) and Belovicx (1979).
- 6) That is, provide the individuals with more than a one-shot survey. Let them answer a survey, report the tentative results of that survey back to them, let them adjust their answers, report the new tentative results, and so forth until an unannounced stopping time. At this stopping time allow the final results to take effect.
- 7) Loeb (1977) considers the general comparability problems associated with relating private good auction mechanisms and public good auction mechanisms.
- 8) Schoemaker (1982) concludes: "As a descriptive model seeking insight into how decisions are made, expected utility theory fails on three counts. First, people do not structure problems as holistically and comprehensively as expected utility theory suggests. Second, they do not process information, especially probabilities, according to the expected utility rule. Finally, expected utility theory, as an "as if" model, poorly predicts choice behavior in laboratory situations. Hence, it is doubtful that expected utility theory should or could serve as a general descriptive model." (p. 552)
- 9) This function is generally a step function. The assumption that individual 1 has a lower Y than individual 2 and so forth is only a simplifying assumption to keep the mathematics simple.
- 10) If individual i maximizes $U_i(Y_i + m_i, 1)$ then he will select an infinite value for m_i . Only a preference for fairness or equity not modelled in this problem would cause m_i to be bounded.